

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2002, 09:20:34 ; Search time 29.93 Seconds
(without alignments)
29.689 Million cell updates/sec

Title: US-09-461-061A-1
Perfect score: 45
Sequence: 1 NNATFYFK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues 747574
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	9	21	Human kininogen D3
2	45	100.0	10	21	Anti-angiogenic D3
3	45	100.0	16	21	Anti-angiogenic D3
4	45	100.0	16	21	Anti-angiogenic D3
5	45	100.0	32	21	Anti-angiogenic D3
6	45	100.0	117	14	Domaine 3, bradyki
7	45	100.0	122	21	Human kininogen D3
8	45	100.0	123	21	Human high mol.wt.
9	45	100.0	248	22	Novel human diagno
10	45	100.0	369	22	Novel human diagno
11	45	100.0	644	22	Novel human diagno

12	39	86.7	26	18	AAW54336	Bradykinin analogo
13	36	80.0	78	22	AAO505080	Human polypeptide
14	34	75.6	42	22	AAG74378	Human colon cancer
15	34	75.6	376	21	AAB28404	Staphylococcus hom
16	34	75.6	541	22	ABBS2487	Escherichia coli p
17	34	75.6	602	20	AAW93820	Bacillus sp. GUS p
18	34	75.6	602	20	AAW93822	Bacillus sp. GUS p
19	34	75.6	602	20	AAW93825	Bacillus sp. codon
20	34	75.6	602	20	AAW93826	Bacillus sp. GUS p
21	34	75.6	602	21	AAB28402	Staphylococcus bet
22	34	75.6	615	21	AAB28408	Codon-optimised St
23	34	75.6	618	20	AAW93821	Bacillus sp. GUS p
24	33	73.3	71	22	AAW82831	Human immune/haema
25	33	73.3	173	21	AAG44316	Arabidopsis thalia
26	33	73.3	251	21	AAG44315	Arabidopsis thalia
27	33	73.3	259	21	AAG44314	Arabidopsis thalia
28	32	71.1	1191	22	ABB60775	Drosophila melanog
29	31	68.9	67	21	AAB34049	Human secreted pro
30	31	68.9	77	22	AAW84819	Human immune/haema
31	31	68.9	89	21	AAG01238	Human secreted pro
32	31	68.9	121	21	AAW81200	Human mutant cysta
33	31	68.9	128	21	AAW81189	Human mutant cysta
34	31	68.9	149	20	AAW36862	Protein which is s
35	31	68.9	161	22	AAO12964	Human polypeptide
36	31	68.9	178	22	ABG09760	Novel human diagno
37	31	68.9	193	16	AAW71919	Macaque platelet a
38	31	68.9	193	18	AAW26505	Macaque platelet-a
39	31	68.9	193	18	AAW23800	Macaque partial pl
40	31	68.9	193	18	AAW09812	Monkey macaque plat
41	31	68.9	193	19	AAW38365	Internal PAF-AH pe
42	31	68.9	193	20	AAW96340	Macaque PAF-AH pe
43	31	68.9	193	20	AAW73365	A macaque platelet
44	31	68.9	193	21	AAW07996	Macaque platelet
45	31	68.9	193	21	AAW88310	Macaque PAF-AH ami

ALIGNMENTS

RESULT 1	
AAW37455	
ID	AAW37455 standard; peptide; 9 AA.
AC	AAW37455;
DT	21-FEB-2001 (first entry)
DE	Human kininogen D3 peptide fragment.
KW	Enzyme; legumain; endopeptidase; cystatin; human; kininogen.
OS	Homo sapiens.
PN	WO200064945-A1.
PD	02-NOV-2000.
PF	20-APR-2000; 2000WO-GB01571.
PR	22-APR-1999; 99GB-0009133.
XX	(BABR-) BABRAHAM INST.
XX	Abrahamson M, Barrett AJ;
XX	WPI; 2000-687316/67.
XX	Inhibition of mammalian legumain or legumain-related endopeptidase by
XX	cystatin involves interaction with second papain-non-reactive site of
XX	cystatin -
XX	Disclosure: Fig 4; 45pp; English.

CC The present invention relates to inhibition of the enzymatic activity of
 CC legumain or a legumain-related endopeptidase by cystatin. The inhibition
 CC involves an interaction between legumain and a papain-non-reactive site
 CC of cystatin. Legumain (EC 3.4.22.34) is a cysteine endopeptidase, and
 CC performs a protein-processing function. The present sequence is a peptide
 CC fragment of human kininogen D3, which was used in the present invention.
 CC Kininogen is a type 3 cystatin. The present sequence is thought to be
 CC involved in a legumain-inhibitory site.

XX Sequence 9 AA;

SQ Query Match 100.0%; Score 45; DB 21; Length 9;

Best Local Similarity 100.0%; Pred. No. 6.4e+05;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNATFFYFK 8
 Db 1 nnatfyfk 8

RESULT 2

AAV95405
 ID AAV95405 standard; Peptide; 10 AA.

XX AC AAV95405;

XX 25-SEP-2000 (first entry)

XX Anti-angiogenic D3 peptide.

XX Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
 XX endothelial cell proliferation; apoptosis; cancer; ocular disorder;
 XX rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
 XX therapy; human; D3 peptide.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Misc-difference 1 /note= "Xaa represents 0 amino acids, or the
 XX sequence of AAV95406 or its N-terminal
 XX truncation fragment containing at least
 XX 1 amino acid"

XX Peptide 2..9

XX /note= "corresponds to residues Asn(275)-Lys(282)
 XX of HK domain 3"

XX Misc-difference 10

XX /note= "Xaa represents 0 amino acids, or the
 XX sequence of AAV95407 or its C-terminal
 XX truncation fragment containing at least
 XX 1 amino acid"

XX WO200035407-A2.

XX 22-JUN-2000.

XX 02-DEC-1999; 99WO-US28465.

XX 16-DEC-1998; 98US-0112427.

XX (UTEM) UNIV TEMPLE.

XX (MCCR/) MCCRAE R K.

XX McCrae RK;

XX WPI; 2000-442247/38.

XX Composition for inhibiting angiogenesis and endothelial cell
 XX proliferation, inducing endothelial cell apoptosis and treating cancer,
 XX rheumatoid arthritis, and ocular disorders comprises a kininogen domain
 XX 3 analog

PS Claim 1; Page 25; 44pp; English.

XX The present sequence is that of a D3 peptide derived from high
 CC mol.wt. kininogen (HK) domain 3 (see AAV95426). The D3 peptide, which
 CC may optionally include N-terminal and/or C-terminal protecting
 CC groups, inhibits endothelial cell proliferation and thus possesses
 CC anti-angiogenic activity. It is an example of peptides of the
 CC invention (see AAV95405-26) which are analogues of certain sites in
 CC the HK domain 3, in this case amino acids Asn275-Lys282. The
 CC peptides inhibit endothelial cell proliferation and may also induce
 CC endothelial cell apoptosis. Compositions including such peptides
 CC are used in claimed methods for inhibiting angiogenesis, inhibiting
 CC endothelial cell proliferation, and inducing endothelial cell
 CC apoptosis. Cancer, rheumatoid arthritis, and ocular disorders
 CC characterized by undesired vascularization of the retina are treated.

XX Sequence 10 AA;

Query Match 100.0%; Score 45; DB 21; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.051;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNATFFYFK 8

Db 2 nnatfyfk 9

RESULT 3

AAV95409

ID AAV95409 standard; Peptide; 16 AA.

XX AC AAV95409;

XX 25-SEP-2000 (first entry)

XX Anti-angiogenic D3 peptide.

XX Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
 XX endothelial cell proliferation; apoptosis; cancer; ocular disorder;
 XX rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
 XX therapy; human; D3 peptide.

XX Homo sapiens.

XX WO200035407-A2.

XX 22-JUN-2000.

XX 02-DEC-1999; 99WO-US28465.

XX 16-DEC-1998; 98US-0112427.

XX (UTEM) UNIV TEMPLE.

XX (MCCR/) MCCRAE R K.

XX McCrae RK;

XX WPI; 2000-442247/38.

XX Composition for inhibiting angiogenesis and endothelial cell
 XX proliferation, inducing endothelial cell apoptosis and treating cancer,
 XX rheumatoid arthritis, and ocular disorders comprises a kininogen domain
 XX 3 analog

PS Claim 6; Page 26; 44pp; English.

XX The present sequence is that of a D3 peptide derived from human
 CC high mol.wt. kininogen (HK) domain 3 (see AAV95426). The D3 peptide
 CC inhibits endothelial cell proliferation and thus possesses
 CC anti-angiogenic activity. It is an example of D3 peptides of the
 CC invention (see AAV95405-26) that are analogues of certain sites in
 CC the HK domain 3, in this case amino acid residues Asn275-Lys282.

CC The peptides inhibit endothelial cell proliferation and may also
 CC induce endothelial cell apoptosis. Compositions including the
 CC peptides are used in claimed methods for inhibiting angiogenesis,
 CC inhibiting endothelial cell proliferation, and inducing endothelial
 CC cell apoptosis. Cancer, rheumatoid arthritis, and ocular disorders
 CC characterized by undesired vascularization of the retina are treated.
 CC The IC50 value for the present peptide was less than 0.8 μ M for
 CC inhibition of fibroblast growth factor-induced HUVEC cell
 CC proliferation.
 CC Sequence 16 AA;

Query Match 100.0%; Score 45; DB 21; Length 16;
 Best Local Similarity 100.0%; Pred. No. 0.082;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNATFYFK 8
 DB 9 nnatfyfk 16

RESULT 4
 AAY95410
 ID AAY95410 standard; Peptide; 16 AA.

XX AC AAY95410;
 XX DT 25-SEP-2000 (first entry)
 XX DE Anti-angiogenic D3 peptide.
 XX KW Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
 KW endothelial cell proliferation; apoptosis; cancer; ocular disorder;
 KW rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
 KW therapy; human; D3 peptide.

XX OS Homo sapiens.

XX PN WO200035407-A2.

XX PD 22-JUN-2000.

XX PF 02-DEC-1999; 99WO-US28465.

XX PR 16-DEC-1998; 98US-0112427.

XX PA (UTEM) UNIV TEMPLE.
 (MCCR/) MCCRAE R K.

XX PI McCrae RK;

XX DR WPI; 2000-442247/38.

XX Composition for inhibiting angiogenesis and endothelial cell
 PT proliferation, inducing endothelial cell apoptosis and treating cancer,
 PT rheumatoid arthritis, and ocular disorders comprises a kininogen domain
 PT 3 analog -

XX Claim 7; Page 26; 44pp; English.

XX The present sequence is that of a D3 peptide derived from human
 CC high mol.wt. kininogen (HK) domain 3 (see AAY95426). The D3 peptide
 CC inhibits endothelial cell proliferation and thus possesses
 CC anti-angiogenic activity. It is an example of D3 peptides of the
 CC invention (see AAY95405-26) that are analogues of certain sites in
 CC the HK domain 3, in this case amino acid residues Asn275-Lys282.
 CC The peptides inhibit endothelial cell proliferation and may also
 CC induce endothelial cell apoptosis. Compositions including the
 CC peptides are used in claimed methods for inhibiting angiogenesis,
 CC inhibiting endothelial cell proliferation, and inducing endothelial
 CC cell apoptosis. Cancer, rheumatoid arthritis, and ocular disorders
 CC characterized by undesired vascularization of the retina are treated.

CC The IC50 value for the present peptide was less than 0.8 μ M for
 CC inhibition of fibroblast growth factor-induced HUVEC cell
 CC proliferation.
 CC Sequence 16 AA;

Query Match 100.0%; Score 45; DB 21; Length 16;
 Best Local Similarity 100.0%; Pred. No. 0.082;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNATFYFK 8
 DB 1 nnatfyfk 8

RESULT 5
 AAY95408
 ID AAY95408 standard; Peptide; 32 AA.

XX AC AAY95408;

XX DT 25-SEP-2000 (first entry)

XX DE Anti-angiogenic D3 peptide.

XX KW Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;
 KW endothelial cell proliferation; apoptosis; cancer; ocular disorder;
 KW rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;
 KW therapy; human; D3 peptide.

XX OS Homo sapiens.

XX PN WO200035407-A2.

XX PD 22-JUN-2000.

XX PF 02-DEC-1999; 99WO-US28465.

XX PR 16-DEC-1998; 98US-0112427.

XX PA (UTEM) UNIV TEMPLE.
 (MCCR/) MCCRAE R K.

XX PI McCrae RK;

XX DR WPI; 2000-442247/38.

XX Composition for inhibiting angiogenesis and endothelial cell
 PT proliferation, inducing endothelial cell apoptosis and treating cancer,
 PT rheumatoid arthritis, and ocular disorders comprises a kininogen domain
 PT 3 analog -

XX Claim 4; Page 26; 44pp; English.

XX The present sequence is that of a D3 peptide derived from human
 CC high mol.wt. kininogen (HK) domain 3 (see AAY95426). The D3 peptide
 CC inhibits endothelial cell proliferation and thus possesses
 CC anti-angiogenic activity. It is an example of D3 peptides of the
 CC invention (see AAY95405-26) that are analogues of certain sites in
 CC the HK domain 3, in this case amino acid residues Asn275-Lys282.
 CC The peptides inhibit endothelial cell proliferation and may also
 CC induce endothelial cell apoptosis. Compositions including the
 CC peptides are used in claimed methods for inhibiting angiogenesis,
 CC inhibiting endothelial cell proliferation, and inducing endothelial
 CC cell apoptosis. Cancer, rheumatoid arthritis, and ocular disorders
 CC characterized by undesired vascularization of the retina are treated.

XX Sequence 32 AA;

Query Match 100.0%; Score 45; DB 21; Length 32;
 Best Local Similarity 100.0%; Pred. No. 0.17;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNATFYFK 8

Db 13 nnatfyfk 20

RESULT 6

AAR33350

ID AAR33350 standard; protein; 117 AA.

XX AC AAR33350;

XX DT 01-JUL-1993 (first entry)

XX DE Domaine 3, bradykinin release activating peptide.

XX KW Domain 3; human; kininogen; heavy chain; low molecular weight; plasma;
KW trypsin; platelet; activation; granule contents; hemostasis; thrombin;
KW tissue plasminogen activator; thrombosis; inflammatory response;
KW endothelial cell; von Willebrand factor;

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide 1..18

FT Protein /note= "Leader peptide"

FT /note= "Mature protein"

XX PN WO9303748-A.

XX PD 04-MAR-1993.

XX PF 13-AUG-1992; 92WO-US06809.

XX PR 13-AUG-1991; 91US-0744545.

XX PA (UTEM) UNIV TEMPLE.

XX PI Jiang Y, Schmaier AB;

XX DR WPI; 1993-093714/11.

XX PT Use of trypsin-cleavage fragment of human kininogen - for
PT increasing vascular bradykinin release, for lowering blood
PT pressure and treating hypertension

XX PS Disclosure; Fig 1; 46pp; English.

XX CC The sequence given represents domain 3, amino acids 246-362, of
CC the human kininogen heavy chain. Domain 3 was isolated from low
CC molecular weight kininogen, derived from human plasma, by cleavage
CC with trypsin. Domain 3 peptide inhibits platelet activation causing
CC a marked decrease in the platelets ability to aggregate and secrete
CC their granule contents. The granule contents comprise proteins which
CC participate in hemostasis, thrombosis and the inflammatory response.
CC Domain 3 also inhibits endothelial cell activation shown by a decrease
CC in secretion of endothelial cell contents such as tissue plasminogen
CC activator and von Willebrand factor. Domain 3 functions to inhibit
CC cell activation by blocking thrombin binding to its target cells, the
CC peptide is a selective inhibitor of thrombin-induced platelet
CC activation.

XX SQ Sequence 117 AA;

Query Match

Best Local Similarity 100.0%; Score 45; DB 14; Length 117;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNATFYFK 8

Db 13 nnatfyfk 20

Db 30 nnatfyfk 37

RESULT 7

AAB37447

ID AAB37447 standard; protein; 122 AA.

XX AC AAB37447;

XX DT 21-FEB-2001 (first entry)

XX DE Human kininogen D3.

XX KW Enzyme; legumain; endopeptidase; cystatin; human; kininogen.

XX OS Homo sapiens.

XX PN WO200064945-A1.

XX PD 02-NOV-2000.

XX PF 20-APR-2000; 2000WO-GB01571.

XX PR 22-APR-1999; 99GB-0009133.

XX PA (BABR-) BABRAHAM INST.

XX PI Abrahamson M, Barrett AJ;

XX DR WPI; 2000-687316/67.

XX PT Inhibition of mammalian legumain or legumain-related endopeptidase by
PT cystatin involves interaction with second papain-non-reactive site of
PT cystatin

XX PS Disclosure; Fig 4; 45pp; English.

XX CC The present invention relates to inhibition of the enzymatic activity of
CC legumain or a legumain-related endopeptidase by cystatin. The inhibition
CC involves an interaction between legumain and a papain-non-reactive site
CC of cystatin. Legumain (EC 3.4.22.34) is a cysteine endopeptidase, and
CC performs a protein-processing function. The present sequence is human
CC kininogen D3, which was used in the present invention. Kininogen is a
CC type 3 cystatin.

XX SQ Sequence 122 AA;

Query Match 100.0%; Score 45; DB 21; Length 122;

Best Local Similarity 100.0%; Pred. No. 0.67;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNATFYFK 8

Db 35 nnatfyfk 42

RESULT 8

AA95426

ID AA95426 standard; Peptide; 123 AA.

XX AC AA95426;

XX DT 25-SEP-2000 (first entry)

XX DE Human high mol.wt. kininogen domain 3.

XX KW Anti-angiogenic; angiogenesis; inhibitor; kininogen; homologue;

XX endothelial cell proliferation; apoptosis; cancer; ocular disorder;

XX KW rheumatoid arthritis; cytostatic; antiarthritic; antirheumatic;

XX KW therapy; human; D3 peptide.

XX OS Homo sapiens.

XX PN WO200035407-A2.
XX PD 22-JUN-2000.
XX PF 02-DEC-1999; 99WO-US28465.
XX PR 16-DEC-1998; 98US-0112427.
XX PA (UTEM) UNIV TEMPLE.
XX PA (MCCR/) MCCRAE R K.
XX PI McCrae RK;
XX DR WPI; 2000-442247/38.
XX PT Composition for inhibiting angiogenesis and endothelial cell
PT proliferation, inducing endothelial cell apoptosis and treating cancer,
PT 3 analog -
XX PS Disclosure; Page 4; 44pp; English.
XX CC The present sequence is that of domain 3 of human high mol.wt.
CC kininogen (HK). The invention provides peptides (see AAY95405-24)
CC that are analogues of certain sites in the HK domain 3,
CC specifically Asn275-Lys282, Cys246-Cys249, Leu331-Tyr338 and
CC Tyr299-Ser314. The peptides, in which native Cys residues may be
CC replaced by Ala residues, inhibit endothelial cell proliferation
CC and may also induce endothelial cell apoptosis. Compositions
CC including the peptides are used in claimed methods for inhibiting
CC angiogenesis, inhibiting endothelial cell proliferation, and
CC inducing endothelial cell apoptosis. Cancer, rheumatoid arthritis,
CC and ocular disorders characterized by undesired vascularization of
CC the retina are treated.
XX SQ Sequence 123 AA;

Query Match 100.0%; Score 45; DB 21; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNATFYFK 8
|||||||
Db 41 nnatfyfk 48

RESULT 9
ABG21102
ID ABG21102 standard; Protein; 248 AA.
XX AC ABG21102;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #21093.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX

XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR N-PSDB; AAS85289.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX PS Claim 20; SEQ ID No 51461; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 248 AA;

Query Match 100.0%; Score 45; DB 22; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNATFYFK 8
|||||||
Db 90 nnatfyfk 97

RESULT 10
ABG21099
ID ABG21099 standard; Protein; 369 AA.
XX AC ABG21099;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #21090.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX

PI Drmanac RT, Liu C, Tang YT;
 XX WPI: 2001-639362/73.
 DR N-PSDB; AAS85286.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 20; SEQ ID No 51458; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 369 AA;

Query Match 100.0%; Score 45; DB 22; Length 369;
 Best Local Similarity 100.0%; Pred. No. 2.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNATFFFK 8
 |||||
 Db 333 nnatfyfk 340

RESULT 11
 ABG21101
 ID ABG21101 standard; Protein: 644 AA.
 XX
 AC ABG21101;
 DT
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #21092.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 XX 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX Drmanac RT, Liu C, Tang YT;
 PI

XX WPI: 2001-639362/73.
 DR N-PSDB; AAS85288.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 20; SEQ ID No 51460; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 644 AA;

Query Match 100.0%; Score 45; DB 22; Length 644;
 Best Local Similarity 100.0%; Pred. No. 3.7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNATFFFK 8
 |||||
 Db 293 nnatfyfk 300

RESULT 12
 AAW54336
 ID AAW54336 standard; peptide: 26 AA.
 XX
 AC AAW54336;
 XX
 DT 30-JUL-1998 (first entry)
 XX
 DE Bradykinin analogous peptide 19.

XX Inhibition; thrombin-induced platelet; prevention; platelet aggregation;
 KW ADP-induced activation.
 XX
 OS Synthetic.
 XX
 PN WO9641640-A1.
 XX
 PD 27-DEC-1996.
 XX
 PF 07-JUN-1996; 96WO-US09940.
 XX
 PR 09-JUN-1995; 95US-0000096.
 XX
 PA (UNMI) UNIV MICHIGAN.
 XX
 PI Hasan AAK, Schmaier AH;
 XX WPI: 1997-065304/06.
 DR

XX Inhibition of platelet activation and aggregation - by admin. of new
PT or known bradykinin analogues
XX Disclosure; Page 44; 73pp; English.
XX
CC Administration of a peptide or multimer related to bradykinin or other
CC disclosed peptides and multimers can be used for the inhibition of
CC thrombin-induced platelets or other cells. They can also be used for
CC preventing platelet aggregation, or inhibiting ADP-induced activation.
CC This is useful to prevent arterial occlusions arising from coronary
CC thrombosis and stroke.
XX Sequence 26 AA;
SQ

Query Match 86.7%; Score 39; DB 18; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NATFYFK 8
Db 1 natfyfk 7
IIIIII

RESULT 13
AAO05080
ID AAO05080 standard; Protein; 78 AA.
XX AC AAO05080;
XX
XX 06-NOV-2001 (first entry)
XX Human polypeptide SEQ ID NO 18972.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
XX Homo sapiens.
XX WO200164835-A2.
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US04927.
XX
XX 28-FEB-2000; 2000US-0515126.
XX 18-MAY-2000; 2000US-0577409.
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-514838/56.
XX N-PSDB; AAI85011.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing
XX diagnosing and treating e.g. leukaemia, inflammation and immune
XX disorders -
XX
XX Claim 20; SEQ ID NO 18972; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 78 AA;
Query Match 80.0%; Score 36; DB 22; Length 78;
Best Local Similarity 85.7%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNATFYF 7
Db 38 nnttffy 44
IIIIII

RESULT 14
AAG74378
ID AAG74378 standard; Protein; 42 AA.
XX AC AAG74378;
XX
XX 03-SEP-2001 (first entry)
XX Human colon cancer antigen protein SEQ ID NO:5142.
XX
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX colorectal carcinoma.
XX
XX Homo sapiens.
XX WO200122920-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US26524.
XX
XX 29-SEP-1999; 99US-0157137.
XX 03-NOV-1999; 99US-0163280.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Barash SC, Birse CE, Rosen CA;
XX WPI; 2001-235357/24.
XX N-PSDB; AAI33809.
XX
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers -
XX
XX Claim 11; Page 6843; 9803pp; English.
XX
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
XX cancer-associated nucleic acid molecules (N) and proteins (P), where
XX the proteins are collectively known as colon cancer antigens. The colon
XX cancer antigens have cytostatic activity and can be used in gene
XX therapy and vaccine production. N and P may be used in the prevention,
XX diagnosis and treatment of diseases associated with inappropriate P
XX expression. For example, N and P may be used to treat disorders
XX associated with decreased expression by rectifying mutations or deletions
XX in a patient's genome that affect the activity of P by expressing P.
XX inactive proteins or to supplement the patients own production of P.
XX Additionally, N may be used to produce the colon cancer-associated Ps,
XX by inserting the nucleic acids into a host cell and culturing the cell
XX to express the proteins. N and P can be used in the prevention, diagnosis
XX and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
XX and AAB77789 represent sequences used in the exemplification of the
XX present invention.
XX
XX N.B. Pages 666 to 682 and page 7053 of the sequence listing were
XX missing at time of publication, meaning no sequences are present for
XX SEQ ID NO:1027 to 1052, 7921 and 7922.

XX SQ Sequence 42 AA;

Query Match 75.6%; Score 34; DB 22; Length 42;
Best Local Similarity 85.7%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNATFFYF 7
| | | | |
Db 22 nratfyyf 28

RESULT 15

AAB28404
ID AAB28404 standard; Protein; 376 AA.

XX AC AAB28404;

XX DT 26-JAN-2001 (first entry)

XX DE Staphylococcus homini beta-D-glucuronidase.

XX KW Microbial; beta-glucuronidase; GUS; Enterobacter; Salmonella;
KW Pseudomonas; Staphylococcus; Thermotoga; transgenic plant; bioindicator;
KW transgenic insect; marker; glucuronide detoxification.

XX OS Staphylococcus homini.

XX PN WO200055333-A1.

XX PD 21-SEP-2000.

XX PF 16-MAR-2000; 2000WO-US07107.

XX PR 17-MAR-1999; 99US-0270957.

XX PA (CAMB-) CAMBIA BIOSYSTEMS LLC.

XX PI Jefferson RA, Mayer JE;

XX DR WPI; 2000-647075/62.

XX DR N-PSDB; AAA07936.

XX PT Novel microbial beta-glucuronidase genes and gene products used as
PT reporter/effector molecule, as diagnostic tool, in positive selection,
PT to target molecules to specific cells and to detect and track linked
PT genes -

XX Claim 3; Fig 5B; 116pp; English.

XX PS The present sequence is a microbial beta-glucuronidase (GUS)
XX CC protein. GUS genes were obtained from six different genera:
XX CC Enterobacter/Salmonella, Pseudomonas, Salmonella, Staphylococcus and
XX CC Thermotoga. Microbial GUS can be used as a reporter/effector molecule for
XX CC transgenic constructions and in vitro diagnostic applications. It may
XX CC also be used to generate sentinel plants that serve as bioindicators of
XX CC environmental status. It may be used to generate transgenic insects for
XX CC tracking insect populations or to facilitate the development of a
XX CC bioassay for compounds that affect molecules critical for insect
XX CC development (e.g. juvenile hormone). Secreted GUS may also serve as a
XX CC marker for beneficial fungi destined for release into the environment. In
XX CC animal systems, secreted GUS may be used to achieve extracellular
XX CC detoxification of glucuronides (e.g. toxin glucuronide) and to examine
XX CC conjugation patterns of glucuronides. Microbial GUS may also be used in
XX CC traditional medical diagnostic assays, for drug testing, pharmacokinetic
XX CC studies, bioavailability studies, diagnosis of diseases and syndromes,
XX CC following progression of disease or its response to therapy. Microbial
XX CC GUS has increased thermal stability, high turnover number and enzymatic
XX CC activity. It is highly specific for the substrate and water soluble, and
XX CC the substrates are stable.

XX SQ Sequence 376 AA;

Query Match 75.6%; Score 34; DB 21; Length 376;
Best Local Similarity 75.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NNATFFYF 8
| | | | |
Db 57 nnkpfyfk 64

Search completed: July 10, 2002, 09:21:54
Job time: 80 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2002, 09:20:34 ; Search time 13.1 Seconds
(without alignments)
14.916 Million cell updates/sec

Title: US-09-461-061a-1
Perfect score: 45
Sequence: 1 NNATFFVK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Issued_Patents_AA:*
- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
 - 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
 - 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
 - 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
 - 5: /cgn2_6/ptodata/2/1aa/PCRTUS.COMB.pep:*
 - 6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	45	100.0	117	1	US-08-193-114B-1		Sequence 1, Appli
2	45	100.0	117	5	PCT-US92-06809-1		Sequence 1, Appli
3	39	86.7	26	4	US-08-676-242-15		Sequence 15, Appl
4	31	68.9	193	1	US-08-483-140-30		Sequence 30, Appl
5	31	68.9	193	2	US-08-483-938A-36		Sequence 36, Appl
6	31	68.9	441	1	US-08-470-187-8		Sequence 8, Appli
7	31	68.9	441	1	US-08-318-905-8		Sequence 8, Appli
8	31	68.9	441	1	US-08-483-232-8		Sequence 8, Appli
9	31	68.9	441	1	US-08-483-140-8		Sequence 8, Appli
10	31	68.9	441	2	US-08-485-938A-8		Sequence 8, Appli
11	31	68.9	441	2	US-08-910-041-8		Sequence 8, Appli
12	31	68.9	441	3	US-09-328-474-8		Sequence 8, Appli
13	31	68.9	441	3	US-09-100-546-8		Sequence 8, Appli
14	31	68.9	441	4	US-09-010-715-8		Sequence 8, Appli
15	31	68.9	441	4	US-09-577-758-8		Sequence 8, Appli
16	31	68.9	444	1	US-08-483-140-28		Sequence 28, Appl
17	31	68.9	444	2	US-08-485-938A-32		Sequence 32, Appl
18	31	68.9	902	1	US-08-701-846-2		Sequence 2, Appli
19	31	68.9	2710	1	US-08-480-604A-6		Sequence 6, Appli
20	31	68.9	2710	2	US-08-405-496A-6		Sequence 6, Appli
21	31	68.9	2710	4	US-08-915-136-6		Sequence 6, Appli
22	30	66.7	84	4	US-09-227-357-597		Sequence 597, App
23	30	66.7	572	6	5200183-5		Patent No. 5200183
24	30	66.7	602	3	US-08-446-100-1		Sequence 1, Appli
25	30	66.7	602	3	US-08-446-100-2		Sequence 2, Appli
26	30	66.7	602	3	US-08-446-100-3		Sequence 3, Appli
27	30	66.7	602	3	US-08-446-100-4		Sequence 4, Appli

28	30	66.7	602	3	US-08-446-100-5	Sequence 5, Appli
29	30	66.7	602	3	US-08-446-100-6	Sequence 6, Appli
30	30	66.7	602	3	US-08-446-100-7	Sequence 7, Appli
31	30	66.7	602	3	US-08-446-100-8	Sequence 8, Appli
32	30	66.7	602	3	US-08-446-100-9	Sequence 9, Appli
33	30	66.7	602	3	US-08-446-100-10	Sequence 10, Appl
34	30	66.7	602	3	US-08-446-100-11	Sequence 11, Appl
35	30	66.7	602	3	US-08-446-100-12	Sequence 12, Appl
36	30	66.7	602	3	US-08-446-100-13	Sequence 13, Appl
37	30	66.7	602	3	US-08-446-100-14	Sequence 14, Appl
38	30	66.7	602	3	US-08-446-100-15	Sequence 15, Appl
39	30	66.7	602	3	US-08-446-100-16	Sequence 16, Appl
40	30	66.7	602	3	US-08-446-100-17	Sequence 17, Appl
41	30	66.7	602	3	US-08-446-100-18	Sequence 18, Appl
42	30	66.7	602	3	US-08-446-100-24	Sequence 24, Appl
43	30	66.7	602	4	US-09-334-489-3	Sequence 3, Appli
44	30	66.7	602	4	US-09-334-489-4	Sequence 4, Appli
45	30	66.7	602	6	5215909-11	Patent No. 5215909

ALIGNMENTS

RESULT 1
US-08-193-114B-1
: Sequence 1, Application US/08193114B
: Patent No. 5472945
: GENERAL INFORMATION:
: APPLICANT: Schmaier, Alvin H.
: APPLICANT: Jiang, Yongping
: TITLE OF INVENTION: Modulation of Blood
: TITLE OF INVENTION: Pressure and Inhibition of Platelet Activation
: TITLE OF INVENTION: with kininogen Fragment
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Seidel, Gonda, Lavorigna &
: ADDRESSEE: Monaco, P.C.
: STREET: 1800 Two Penn Center Plaza
: CITY: Philadelphia
: STATE: Pennsylvania
: COUNTRY: U.S.A.
: ZIP: 19102
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
: COMPUTER: IBM PS/2
: OPERATING SYSTEM: MS-DOS
: SOFTWARE: WordPerfect 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/193,114B
: FILING DATE: 9 February 1994
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: U.S. Application
: APPLICATION NUMBER: Serial No. 5472945 07/744,545
: FILING DATE: 13 August 1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Monaco, Daniel A.
: REGISTRATION NUMBER: 30,480
: REFERENCE/DOCKET NUMBER: 6056-137 CII
: TELEPHONE: (215) 568-8383
: TELEFAX: (215) 568-5549
: TELEX: No. 5472945e
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 117 amino acids
: TYPE: peptide
: TOPOLOGY: linear
US-08-193-114B-1

SEA 7D No.1

Query Match 100.0%; Score 45; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 0.16;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNATFYFK 8

|||||||

Db 30 NNATFYFK 37

RESULT 2

PCT-US92-06809-1

; Sequence 1, Application PC/TUS9206809

; GENERAL INFORMATION:

; APPLICANT: Schmaier, Alvin H.

; APPLICANT: Jiang, Yongping

; TITLE OF INVENTION: Modulation of Blood

; TITLE OF INVENTION: Pressure by Altering Bradykinin Levels

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Temple University - Of the

; ADDRESSEE: Commonwealth System of Higher Education

; STREET: 406 University Services

; STREET: Building

; CITY: Philadelphia

; STATE: Pennsylvania

; COUNTRY: U.S.A.

; ZIP: 19122

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: Wordperfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US92/06809

; FILING DATE: 19910813

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: U.S. Application

; APPLICATION NUMBER: Serial No. 744,545

; FILING DATE: 13 August 1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Monaco, Daniel A.

; REGISTRATION NUMBER: 30,480

; REFERENCE/DOCKET NUMBER: 6056-137

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (215) 568-8383

; TELEFAX: (215) 568-5549

; TELEX:

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 117 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

PCT-US92-06809-1

Query Match

; Sequence 15, Application US/08676242C

; Best Local Similarity 100.0%; Pred. No. 0.16;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNATFYFK 8

|||||||

Db 30 NNATFYFK 37

RESULT 3

US-08-676-242-15

; Sequence 15, Application US/08676242C

; Patent No. 6143719

; GENERAL INFORMATION:

; APPLICANT: The Regents of the University of Michigan

; APPLICANT: Schmaier, Alvin H.

; APPLICANT: Hasan, Ahmed A.K.

; TITLE OF INVENTION: Bradykinin Analogs As Selective Thrombin Inhibitors

; FILE REFERENCE: 8820-2 US

US-08-483-140-30

; CURRENT APPLICATION NUMBER: US/08/676,242C

; CURRENT FILING DATE: 2000-07-16

; EARLIER APPLICATION NUMBER: 60/000,096

; EARLIER FILING DATE: 1995-06-09

; EARLIER APPLICATION NUMBER: PCT/US96/09940

; EARLIER FILING DATE: 1996-06-07

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 15

; LENGTH: 26

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Bradykinin

; OTHER INFORMATION: analog

US-08-676-242-15

Query Match

; Best Local Similarity 100.0%; Pred. No. 0.43;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NATFYFK 8

|||||||

Db 1 NATFYFK 7

RESULT 4

US-08-483-140-30

; Sequence 30, Application US/08483140

; Patent No. 5698403

; GENERAL INFORMATION:

; APPLICANT: ICOS Corporation

; TITLE OF INVENTION: platelet-Activating Factor Acetyl

; TITLE OF INVENTION: Hydrolase

; NUMBER OF SEQUENCES: 30

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gertein, Murray & Borun

; STREET: 6300 Sears Tower, 233 South Wacker Drive

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/483,140

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/318,905

; FILING DATE: 6-OCT-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/133,803

; FILING DATE: 6-OCT-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: No. 5698403and, Greta E.

; REGISTRATION NUMBER: 35,302

; REFERENCE/DOCKET NUMBER: 32781

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (312) 474-6300

; TELEFAX: (312) 474-0448

; TELEX: 25-3658

; INFORMATION FOR SEQ ID NO: 30:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 193 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

Query Match 68.9%; Score 31; DB 1; Length 193;
Best Local Similarity 71.4%; Pred. No. 89;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 NATYFK 8
: : : : :
Db 143 SATYFK 149

RESULT 5
US-08-485-938A-36
; Sequence 36, Application US/08485938A
; Patent No. 5847088
; GENERAL INFORMATION:
; APPLICANT: Cousins, Lawrence S.
; APPLICANT: Eberhardt, Christine D.
; APPLICANT: Gray, Patrick W.
; APPLICANT: Le Trong, Hai
; APPLICANT: Tjoelker, Larry W.
; APPLICANT: Wilder, Cheryl L.
; TITLE OF INVENTION: Platelet-Activating Factor
; TITLE OF INVENTION: Acetylhydrolase
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,938A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5847088and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/32792
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-0448
; TELEFAX: (312) 474-0448
; TELEX: 25-3658
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 193 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-485-938A-36

Query Match 68.9%; Score 31; DB 2; Length 193;
Best Local Similarity 71.4%; Pred. No. 89;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 NATYFK 8
: : : : :
Db 143 SATYFK 149

RESULT 6
US-08-470-187-8
; Sequence 8, Application US/08470187
; Patent No. 5532152
; GENERAL INFORMATION:
; APPLICANT: Cousins, Lawrence S.
; APPLICANT: Eberhardt, Christine E.
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; APPLICANT: Wilder, Cheryl L.
; TITLE OF INVENTION: Platelet-Activating Factor Acetyl
; TITLE OF INVENTION: Hydrolase
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,187
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5532152and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31672
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3658
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 441 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-470-187-8

Query Match 68.9%; Score 31; DB 1; Length 441;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 NATYFK 8
: : : : :
Db 185 SATYFK 191

RESULT 7
US-08-318-905-8
; Sequence 8, Application US/08318905
; Patent No. 5641669
; GENERAL INFORMATION:
; APPLICANT: Cousins, Lawrence S.
; APPLICANT: Eberhardt, Christine D.
; APPLICANT: Gray, Patrick W.
; APPLICANT: Le Trong, Hai
; APPLICANT: Tjoelker, Larry W.
; APPLICANT: Wilder, Cheryl L.
; TITLE OF INVENTION: Platelet-Activating Factor Acetyl
; TITLE OF INVENTION: Hydrolase
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,905
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/133,803
FILING DATE: 6-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: No. 564169and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 32205
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3658
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 441 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-318-905-8

Query Match 68.9%; Score 31; DB 1; Length 441;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NATYFK 8
Db 185 SATYFK 191

RESULT 8
US-08-483-232-8
Sequence 8, Application US/08483232
Patent No. 5656431
GENERAL INFORMATION:
APPLICANT: Cousens, Lawrence S.
APPLICANT: Eberhardt, Christine D.
APPLICANT: Gray, Patrick W.
APPLICANT: Le Trong, Hai
APPLICANT: Tjoelker, Larry W.
APPLICANT: Wilder, Cheryl L.
TITLE OF INVENTION: Platelet-Activating Factor
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,232
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/318,905
FILING DATE: 6-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/133,803
FILING DATE: 6-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: No. 565643land, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/32689
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3658
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 441 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-483-232-8

Query Match 68.9%; Score 31; DB 1; Length 441;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NATYFK 8
Db 185 SATYFK 191

RESULT 9
US-08-483-140-8
Sequence 8, Application US/08483140
Patent No. 5698403
GENERAL INFORMATION:
APPLICANT: ICOS Corporation
TITLE OF INVENTION: Platelet-Activating Factor Acetyl
TITLE OF INVENTION: Hydrolase
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,140
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/318,905
FILING DATE: 6-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/133,803
FILING DATE: 6-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: No. 5698403and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 32781
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3658
INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 441 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-483-140-8

Query Match 68.9%; Score 31; DB 1; Length 441;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 NATFEYK 8
Db 185 SATYFK 191

RESULT 10
US-08-485-938A-8
; Sequence 8, Application US/08485938A
; Patent No. 5847088
; GENERAL INFORMATION:
; APPLICANT: Cousins, Lawrence S.
; APPLICANT: Eberhardt, Christine D.
; APPLICANT: Gray, Patrick W.
; APPLICANT: Le Trong, Hai
; APPLICANT: Tjoelker, Larry W.
; APPLICANT: Wilder, Cheryl L.
; TITLE OF INVENTION: Platelet-Activating Factor
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,938A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/318,905
; FILING DATE: 06-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/133,803
; FILING DATE: 06-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5847088and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/32792
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3658
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 441 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-485-938A-8

Query Match 68.9%; Score 31; DB 2; Length 441;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 NATFEYK 8
Db 185 SATYFK 191

RESULT 11
US-08-910-041-8
; Sequence 8, Application US/08910041
; Patent No. 5977308
; GENERAL INFORMATION:
; APPLICANT: Cousins, Lawrence S.
; APPLICANT: Eberhardt, Christine D.
; APPLICANT: Gray, Patrick W.
; APPLICANT: Le Trong, Hai
; APPLICANT: Tjoelker, Larry W.
; APPLICANT: Wilder, Cheryl L.
; TITLE OF INVENTION: Platelet-Activating Factor
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,041
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/483,232
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/318,905
; FILING DATE: 06-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/133,803
; FILING DATE: 06-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/34026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3658
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 441 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-910-041-8

Query Match 68.9%; Score 31; DB 2; Length 441;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 NATFEYK 8
Db 185 SATYFK 191

RESULT 12
US-09-328-474-8

; Sequence 8, Application US/09328474
; Patent No. 6045794
; GENERAL INFORMATION:
; APPLICANT: Cousens, Lawrence S.
; APPLICANT: Eberhardt, Christine D.
; APPLICANT: Gray, Patrick W.
; APPLICANT: Le Trong, Hai
; APPLICANT: Tjoelker, Larry W.
; APPLICANT: Wilder, Cheryl L.
; TITLE OF INVENTION: Platelet-Activating Factor
; TITLE OF INVENTION: Acetylhydrolase
; CORRESPONDENCE ADDRESS:
; NUMBER OF SEQUENCES: 30
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/328,474
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/483,232
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/318,905
; FILING DATE: 06-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/133,803
; FILING DATE: 06-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/34026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3658
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 441 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-328-474-8

Query Match 68.9%; Score 31; DB 3; Length 441;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 2 NATFYFK 8
Db 185 SATYFK 191

RESULT 13
US-09-100-546-8
; Sequence 8, Application US/09100546
; Patent No. 6059836
; GENERAL INFORMATION:
; APPLICANT: Cousens, Lawrence S.
; APPLICANT: Eberhardt, Christine D.
; APPLICANT: Gray, Patrick W.
; APPLICANT: Le Trong, Hai
; APPLICANT: Tjoelker, Larry W.

; APPLICANT: Wilder, Cheryl L.
; TITLE OF INVENTION: Platelet-Activating Factor
; TITLE OF INVENTION: Acetylhydrolase
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,546
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/010,715
; FILING DATE:
; APPLICATION NUMBER: US 08/318,905
; FILING DATE: 06-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/133,803
; FILING DATE: 06-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6099836and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/32793
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3658
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 441 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-100-546-8

Query Match 68.9%; Score 31; DB 3; Length 441;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 2 NATFYFK 8
Db 185 SATYFK 191

RESULT 14
US-09-010-715-8
; Sequence 8, Application US/09010715
; Patent No. 6146625
; GENERAL INFORMATION:
; APPLICANT: Cousens, Lawrence S.
; APPLICANT: Eberhardt, Christine D.
; APPLICANT: Gray, Patrick W.
; APPLICANT: Le Trong, Hai
; APPLICANT: Tjoelker, Larry W.
; APPLICANT: Wilder, Cheryl L.
; TITLE OF INVENTION: Platelet-Activating Factor
; TITLE OF INVENTION: Acetylhydrolase
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois

; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/010,715
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/318,905
; FILING DATE: 06-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/133,803
; FILING DATE: 06-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 614662sand, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/32793
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3658
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 441 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-010-715-8

Query Match 68.9%; Score 31; DB 4; Length 441;
Best Local Similarity 71.4%; Pred. NO. 2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 NATVEFK 8
Db 185 SATYVFK 191

RESULT 15
US-09-577-758-8
; Sequence 8, Application US/09577758
; Patent No. 6203790
; GENERAL INFORMATION:
; APPLICANT: Cousens, Lawrence S.
; APPLICANT: Eberhardt, Christine D.
; APPLICANT: Gray, Patrick W.
; APPLICANT: Le Trong, Hai
; APPLICANT: Tjoelker, Larry W.
; APPLICANT: Wilder, Cheryl L.
; TITLE OF INVENTION: Platelet-Activating Factor
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/577,758
; FILING DATE:
; CLASSIFICATION:

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/010,715
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/133,803
; FILING DATE: 06-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6203790and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/32793
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3658
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 441 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-577-758-8

Query Match 68.9%; Score 31; DB 4; Length 441;
Best Local Similarity 71.4%; Pred. NO. 2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 NATVEFK 8
Db 185 SATYVFK 191

Search completed: July 10, 2002, 09:20:56
Job time: 22 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2002, 09:20:34 ; Search time 14.91 Seconds
(without alignments)
51.557 Million cell updates/sec

Title: US-09-461-061a-1
Perfect score: 45
Sequence: 1 NNATFYFK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	427	1 KGHUL1	kininogen, LMW pre
2	45	100.0	644	1 KGHUL1	kininogen, LMW pre
3	39	86.7	433	2 A28055	K-kininogen, LMW I
4	39	86.7	639	2 A25486	kininogen, HMW I p
5	36	80.0	858	2 T24062	hypothetical prote
6	36	80.0	1132	2 T31107	telomerase reverse
7	35	77.8	182	2 T24206	hypothetical prote
8	34	75.6	71	2 B3803	hypothetical prote
9	34	75.6	182	2 T06978	ABA-induced plasma
10	34	75.6	273	2 F86924	hypothetical prote
11	34	75.6	319	2 A86650	rhannosyltransfera
12	34	75.6	440	2 T11319	NADH dehydrogenase
13	34	75.6	497	2 G96611	probable cytochrom
14	34	75.6	630	2 T25830	hypothetical prote
15	34	75.6	758	2 T31994	hypothetical prote
16	34	75.6	949	2 H97322	DNA/RNA helicase,
17	33	73.3	303	2 T32658	hypothetical prote
18	33	73.3	332	2 T01483	hypothetical prote
19	33	73.3	469	2 F69403	hypothetical prote
20	33	73.3	603	2 S70849	cholinesterase (EC
21	33	73.3	706	2 D84466	hypothetical prote
22	33	73.3	895	2 T23191	hypothetical prote
23	32	71.1	484	2 A40774	phosphocholine-bin
24	32	71.1	562	2 A08852	secretory protein
25	32	71.1	563	2 S54420	invasion protein i
26	32	71.1	567	2 E91095	type III secretion
27	32	71.1	567	2 A85941	type III secretion
28	32	71.1	606	2 T10982	NADH dehydrogenase
29	32	71.1	1215	2 T25078	hypothetical prote

30	32	71.1	1659	1 OKBYN2	protein kinase GCN
31	31	68.9	140	2 T10059	cytokinin-induced
32	31	68.9	148	2 B71529	hypothetical prote
33	31	68.9	171	2 F90088	hypothetical prote
34	31	68.9	190	2 F84366	hypothetical prote
35	31	68.9	196	2 F84941	amidotransferase h
36	31	68.9	202	2 AF0604	probable permease
37	31	68.9	211	2 T34501	hypothetical prote
38	31	68.9	240	2 A46179	U2 snRNP auxiliary
39	31	68.9	249	2 F69409	conserved hypothet
40	31	68.9	262	2 S59078	conserved hypothet
41	31	68.9	264	1 JC6125	U2 small nuclear r
42	31	68.9	334	1 RDECEP	N-acetyl-gamma-glu
43	31	68.9	334	2 G91239	N-acetyl-gamma-glu
44	31	68.9	334	2 D86087	N-acetyl-gamma-glu
45	31	68.9	350	2 A82299	outer membrane pro

ALIGNMENTS

RESULT 1

KGHUL1
kininogen, LMW precursor [validated] - human
N:Alternate names: alpha-2-thiol proteinase inhibitor; preprokininogen
N:Contains: bradykinin (kallidin); kininogen I; kininogen II; prokininogen
C:Species: Homo sapiens (man)
C:Date: 06-Jul-1982 #sequence_revision 27-Nov-1985 #text_change 08-Dec-2000
C:Accession: A01280; B25276; A27900; A27699; A31905; A34030
R:Ohkubo, I.; Kurachi, K.; Takasawa, T.; Shiohara, H.; Sasaki, M.
Biochemistry 23, 5691-5697, 1984
A:Title: Isolation of a human cDNA for alpha-2-thiol proteinase inhibitor and its ide
A:Reference number: A90490; MUID:85122621
A:Accession: A01280
A:Molecule type: mRNA
A:Residues: 1-427 <OHK>
A:Cross-references: GB:K02566; NID:g177889; PIDN:AAA35497.1; PID:g177890
R:Takagaki, Y.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 8601-8609, 1985
A:Title: Cloning and sequence analysis of cDNAs for human high molecular weight and 1
A:Reference number: A92544; MUID:85234582
A:Accession: B25276
A:Molecule type: mRNA
A:Residues: 1-427 <TAK>
A:Cross-references: GB:M11437; NID:g186751; PIDN:AAB59551.1; PID:g386853
R:Lottspeich, F.; Kellermann, J.; Henschen, A.; Rauth, G.; Mueller-Esterl, W.
in Kinins IV, part A, Greenbaum, L.M., and Margolius, H.S., eds., pp.91-95, Plenum, N
A:Title: Amino acid sequence of the light chain of human low molecular mass kininogen
A:Reference number: A27900
A:Accession: A27900
A:Molecule type: protein
A:Residues: 390-427 <LOT>
R:Mindroul, T.; Carretero, O.A.; Proud, D.; Walz, D.; Scicli, A.G.
Biochem. Biophys. Res. Commun. 152, 519-526, 1988
A:Title: A new kinin moiety in human plasma kininogens.
A:Reference number: A27699; MUID:88209021
A:Accession: A27699
A:Molecule type: protein
A:Residues: 380-389 <MIN>
R:Maeda, H.; Matsumura, Y.; Kato, H.
J. Biol. Chem. 263, 16051-16054, 1988
A:Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic f
A:Reference number: A31905; MUID:89034061
A:Accession: A31905
A:Molecule type: protein
A:Residues: 381-389 <MAE>
R:Sasaguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.
Biochem. Biophys. Res. Commun. 150, 511-516, 1988
A:Title: Identification of [hydroxyproline(3)]-lysyl-bradykinin released from human p
A:Reference number: A34030; MUID:88106632
A:Accession: A34030
A:Molecule type: protein
A:Residues: 380-389 <SAS>

R:Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.
 J. Biol. Chem. 260, 8610-8617, 1985
 A:Title: Structural organization of the human kininogen gene and a model for its evolution
 A:Reference number: A92545; MUID:85234583
 A:Contents: annotation; gene organization
 R:Pierce, J.V.
 Fed. Proc. 27, 52-57, 1968
 A:Title: Structural features of plasma kinins and kininogens.
 A:Reference number: A91455; MUID:90255622
 A:Contents: annotation; bradykinin
 C:Comment: The LMW kininogen precursor is produced from the same gene as the HMW form (S)
 C:Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the
 C:Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, and
 C:Comment: xiprolin residue is present in the kininogen prior to the release of bradykinin.
 C:Genetics:
 A:Gene: GDB:KNG
 A:Cross-references: GDB:125256; OMIM:228960
 A:Map position: 3q27-3q27
 A:Introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3; 401/3
 C:Superfamily: kininogen; cystatin homology
 C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; glyc
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-427/Product: LMW prokininogen (kininogen I) #status predicted <MAT>
 F:19-389/390-427/Product: LMW kininogen II #status predicted <MAT2>
 F:19-379/Product: LMW kininogen heavy chain #status predicted <HCH>
 F:19-131/Domain: cystatin homology <CX1>
 F:142-253/Domain: cystatin homology <CX2>
 F:264-375/Domain: cystatin homology <CX3>
 F:380-389/Product: lysyl-bradykinin (kallidin II) #status experimental <KB DY>
 F:381-389/Product: bradykinin (kallidin I) #status experimental <BDY>
 F:390-427/Product: LMW kininogen light chain #status experimental <LCH>
 F:19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
 F:28-407,83-94,107-126,142-145,206-218,229-248,264-267,328-340,351-370/Disulfide bonds:
 F:48,169,205,294/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:379-380/Cleavage site: Met-Lys (kallikrein) #status experimental
 F:383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
 F:389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
 F:401/Binding site: carbohydrate (Thr) (covalent) #status absent

Query Match 100.0%; Score 45; DB 1; Length 427;
 Best Local Similarity 100.0%; Pred. No. 0.35;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNATYFK 8
 |||||
 Db 293 NNATYFK 300

RESULT 2
 KGHUHI
 N:Alternate names: alpha-2-thiol proteinase inhibitor; prokininogen; prokininogen
 N:Contains: bradykinin (kallidin I); HMW kininogen I; HMW kininogen II; low molecular we
 C:Species: Homo sapiens (man)
 C:Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 08-Dec-2000
 C:Accession: A01279; A25276; S32422; A91153; A24871; A27699; A31905; A34030; S02
 R:Ohkubo, I.; Kurachi, K.; Takasawa, T.; Shiokawa, H.; Sasaki, M.
 Biochemistry 23, 5691-5697, 1984
 A:Title: Isolation of a human cDNA for alpha-2-thiol proteinase inhibitor and its ident
 A:Reference number: A90490; MUID:85122621
 A:Accession: A01279
 A:Molecule type: mRNA
 A:Residues: 1-389 <OHK>
 A:Cross-references: GB:K02566; NID:g177889
 R:Tagakaki, Y.; Kitamura, N.; Nakanishi, S.
 J. Biol. Chem. 260, 8601-8609, 1985
 A:Title: Cloning and sequence analysis of cDNAs for human high molecular weight and low
 A:Reference number: A92544; MUID:85234582
 A:Accession: A25276
 A:Molecule type: mRNA
 A:Residues: 1-592, 1', 594-644 <TK>
 A:Cross-references: GB:M11437; NID:g186751; PIDN:AA859550.1; PID:g386852

R:Auerswald, E.A.; Roessler, D.; Mentele, R.; Assfalg-Machleidt, I.
 FEBS Lett. 321, 93-97, 1993
 A:Title: Cloning, expression and characterization of human kininogen domain 3.
 A:Reference number: S32422; MUID:92223854
 A:Accession: S32422
 A:Molecule type: mRNA
 A:Residues: 'ANSM', 253-377 <AUE>
 A:Note: differences are due to known cloning artifacts
 R:Lottspeich, F.; Kellermann, J.; Henschen, A.; Foertsch, B.; Muller-Esterl, W.
 Eur. J. Biochem. 152, 307-314, 1985
 A:Title: The amino acid sequence of the light chain of human high-molecular-mass kini
 A:Reference number: A91153; MUID:86030270
 A:Accession: A91153
 A:Molecule type: protein
 A:Residues: 379-644 <LOT>
 A:Note: the bradykinin sequence preceding the light chain sequence was not determined
 R:Kellermann, J.; Lottspeich, F.; Henschen, A.; Mueller-Esterl, W.
 Eur. J. Biochem. 154, 471-478, 1986
 A:Title: Completion of the primary structure of human high-molecular-mass kininogen.
 A:Reference number: A24871; MUID:86108361
 A:Accession: A24871
 A:Molecule type: protein
 A:Residues: 'Z', 20-380 <KEL1>
 R:Kellermann, J.; Lottspeich, F.; Henschen, A.; Mueller-Esterl, W.
 in Kinins IV, Greenbaum, L.M., and Margolius, H.S., ed., pp.85-89, Plenum Press, New
 A:Title: Amino acid sequence of the light chain of human high molecular mass kininoge
 A:Reference number: A27899
 A:Accession: A27899
 A:Molecule type: protein
 A:Residues: 379-389, 'K', 390-407, 'Q', 409-644 <KEL2>
 R:Mindroiu, T.; Carretero, O.A.; Proud, D.; Walz, D.; Scicli, A.G.
 Biochem. Biophys. Res. Commun. 152, 519-526, 1988
 A:Title: A new kinin moiety in human plasma kininogens.
 A:Reference number: A27699; MUID:88205021
 A:Accession: A27699
 A:Molecule type: protein
 A:Residues: 380-389 <MIN>
 R:Maeda, H.; Matsumura, Y.; Kato, H.
 J. Biol. Chem. 263, 16051-16054, 1988
 A:Title: Purification and identification of [hydroxyprolyl(3)]bradykinin in ascitic f
 A:Reference number: A31905; MUID:89034061
 A:Accession: A31905
 A:Molecule type: protein
 A:Residues: 381-389 <MAE>
 R:Sasaguri, M.; Ikeda, M.; Ideishi, M.; Arakawa, K.
 Biochem. Biophys. Res. Commun. 150, 511-516, 1988
 A:Title: Identification of [hydroxyproline(3)]-lysyl-bradykinin released from human p
 A:Reference number: A34030; MUID:88106632
 A:Accession: A34030
 A:Molecule type: protein
 A:Residues: 380-389 <SAS>
 R:Lenarcic, B.; Gabrijelcic, D.; Rozman, B.; Drobnic-Kosorok, M.; Turk, V.
 Biol. Chem. Hoppe-Seyler 369, 257-261, 1988
 A:Title: Human cathepsin B and cysteine proteinase inhibitors (CPIs) in inflammatory
 A:Reference number: S02482; MUID:89076517
 A:Accession: S02482
 A:Molecule type: protein
 A:Residues: 1-19,189-192; 310-314; 381-389 <LEN1>
 R:Kato, H.; Matsumura, Y.; Maeda, H.
 FEBS Lett. 232, 252-254, 1988
 A:Title: Isolation and identification of hydroxyproline analogues of bradykinin in hu
 A:Reference number: A61495; MUID:88211869
 A:Accession: A61495
 A:Molecule type: protein
 A:Residues: 380-389 <KAT1>
 A:Experimental source: urine
 A:Note: this peptide had Pro-383 modified to 4-hydroxyproline
 A:Accession: B61495
 A:Molecule type: protein
 A:Residues: 381-389 <KAT2>
 A:Experimental source: urine
 A:Note: this peptide had Pro-383 modified to 4-hydroxyproline
 A:Accession: C61495

A:Molecule type: protein
A:Residues: 380-389 <KAT3>
R:Lenarcic, B.; Krasovec, M.; Ritonja, A.; Olafsson, I.; Turk, V.
FEBS Lett. 280, 211-215, 1991
A:Title: Inactivation of human cystatin C and kininogen by human cathepsin D.
A:Reference number: S14303; MUID:91192133
A:Accession: S14447
A:Molecule type: protein
A:Residues: 264-359, 'N', 361-375 <LEN2>
R:Little, S.S.; Johnson, D.A.
Biochem. J. 307, 341-346, 1995
A:Title: Human mast cell tryptase isoforms: separation and examination of substrate-specificity
A:Reference number: S55239; MUID:95251593
A:Accession: S55239
A:Molecule type: protein
A:Residues: 450-452, 'X', 454, 'X', 456 <LIT>
R:Straczek, J.; Maechli, F.; le Nguyen, D.; Becchi, M.; Heulin, M.H.; Nabet, P.; Belleville, J.
FEBS Lett. 373, 207-211, 1995
A:Title: Purification from human plasma of a tetrapeptide that potentiates insulin-like growth factor release
A:Reference number: S68059; MUID:96033974
A:Accession: S68059
A:Molecule type: protein
A:Residues: 431-434 <STR>
R:Kitamura, N.; Kitagawa, H.; Fukushima, D.; Takagaki, Y.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 260, 8610-8617, 1985
A:Title: Structural organization of the human kininogen gene and a model for its evolution
A:Reference number: A92545; MUID:85234583
A:Accession: A92545
A:Contents: annotation; gene organization
R:Pierce, J.V.
Fed. Proc. 27, 52-57, 1968
A:Title: Structural features of plasma kinins and kininogens.
A:Reference number: A91455; MUID:90255622
A:Accession: A91455
A:Contents: annotation; bradykinin
C:Comment: The HMW kininogen precursor and the LMW form are produced from the same gene.
C:Comment: Kininogen is a cysteine proteinase inhibitor, takes part in initiation of the blood coagulation cascade.
C:Comment: The glycine/histidine/lysine-rich region of HMW kininogen light chain is important for the release of bradykinin.
C:Comment: Bradykinin, released from kininogen by kallikrein, is a potent vasodilator, and its release is regulated by the presence of the kininogen prior to the release of bradykinin.
C:Genetics:
A:Gene: GDB:KNG
A:Cross-references: GDB:125256; OMIM:228960
A:Map position: 3q27-3q28
A:Introns: 65/3; 102/3; 131/1; 188/3; 224/3; 253/1; 310/3; 346/3; 375/3
C:Superfamily: Kininogen; cystatin homology
C:Keywords: alternative splicing; blood coagulation; cysteine proteinase inhibitor; duplication
F:1-18/Domain: signal sequence #status experimental <SIG>
F:19-644/Product: HMW kininogen I (prokininogen) #status experimental <MAT1>
F:19-379,390-644/Product: HMW kininogen II #status experimental <MAT2>
F:19-379/Domain: HMW kininogen heavy chain #status experimental <HCH>
F:19-131/Domain: cystatin homology <CY1>
F:142-253/Domain: cystatin homology <CY2>
F:264-375/Domain: cystatin homology <CY3>
F:380-389/Product: lysyl-bradykinin (kallidin II) #status experimental <KBDY>
F:381-389/Product: bradykinin (kallidin I) #status experimental <BDY>
F:390-644/Domain: HMW kininogen light chain #status experimental <LCH>
F:421-510/Region: glycine/histidine/lysine-rich 30-residue repeats
F:431-434/Product: low molecular weight growth promoting factor #status experimental <GF>
F:19/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimental
F:28-614,83-94,107-126,142-145,206-218,229-248,264-267,328-340,351-370/Disulfide bonds:
F:48/Binding site: carboxylate (Asn) (covalent) #status experimental
F:169,205,294/Binding site: carboxylate (Asn) (covalent) #status experimental
F:379-380/Cleavage site: Met-Lys (kallikrein) #status experimental
F:383/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
F:389-390/Cleavage site: Arg-Ser (kallikrein) #status experimental
F:401,533,542,546,557,571,593,628/Binding site: carboxylate (Thr) (covalent) #status experimental
F:577/Binding site: carboxylate (Ser) (covalent) #status experimental

Query Match 100.0%; Score 45; DB 1; Length 644;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNATFEYFK 8

Db 293 NNATFEYFK 300

RESULT 3

A28055
K-kininogen, LMW I precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 20-Jun-1989 #sequence_revision 20-Jun-1989 #text_change 15-Nov-1996
C:Accession: A28055
R:Furuto-Kato, S.; Matsumoto, A.; Kitamura, N.; Nakanishi, S.
J. Biol. Chem. 260, 12054-12059, 1985
A:Title: Primary structures of the mRNAs encoding the rat precursors for bradykinin a and b.
A:Reference number: A92496; MUID:86008264
A:Accession: A28055
A:Molecule type: mRNA
A:Residues: 1-433 <FUR>
C:Superfamily: Kininogen; cystatin homology
C:Keywords: alternative splicing
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-433/Product: k-kininogen, LMW I #status predicted <MAT>
F:19-131/Domain: cystatin homology <CY1>
F:142-253/Domain: cystatin homology <CY2>
F:264-375/Domain: cystatin homology <CY3>

Query Match 86.7%; Score 39; DB 2; Length 433;
Best Local Similarity 87.5%; Pred. No. 5.1;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNATFEYFK 8

Db 293 NNATFEYFK 300

RESULT 4

A25486
kininogen, HMW I precursor - rat
N:Contains: bradykinin
C:Species: Rattus norvegicus (Norway rat)
C:Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 15-Nov-1996
C:Accession: A25486
R:Kitagawa, H.; Kitamura, N.; Hayashida, H.; Miyata, T.; Nakanishi, S.
J. Biol. Chem. 262, 2190-2198, 1987
A:Title: Differing expression patterns and evolution of the rat kininogen gene family
A:Reference number: A92625; MUID:87137443
A:Accession: A25486
A:Molecule type: mRNA
A:Residues: 1-639 <KIT>
A:Note: the authors translated the codon CAA for residue 347 as Asn
C:Superfamily: Kininogen; cystatin homology
C:Keywords: alternative splicing
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-639/Product: kininogen, HMW I #status predicted <MAT>
F:19-131/Domain: cystatin homology <CY1>
F:142-253/Domain: cystatin homology <CY2>
F:264-375/Domain: cystatin homology <CY3>

Query Match 86.7%; Score 39; DB 2; Length 639;
Best Local Similarity 87.5%; Pred. No. 7.5;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNATFEYFK 8

Db 293 NNATFEYFK 300

RESULT 5

T24062
hypotheical protein R09A8.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T24062
R:Wilkinson, J.
Submitted to the EMBL Data Library, November 1995
A:Reference number: Z19836
A:Accession: T24062
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-858 <WIL>
A:Cross-references: EMBL:Z68009; PIDN:CAA92004.1; GSPDB:GN00028; CESP:R09A8.2
A:Experimental source: clone R09A8
C:Genetics:
A:Gene: CESP:R09A8.2
A:Map position: X
A:Introns: 197/1; 324/3; 377/3; 435/3; 495/3; 587/1; 627/3; 730/3; 812/1

Query Match 80.0%; Score 36; DB 2; Length 858;
Best Local Similarity 75.0%; Pred. No. 38;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NNATFYFK 8
| :|||||
Db 49 NESTFYFK 56

RESULT 6
T31107
telomerase reverse transcriptase - Oxytricha trifallax
C:Species: Oxytricha trifallax
C>Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
C:Accession: T31107
R:Bryan, T.M.; Sperger, J.M.; Chapman, K.B.; Cech, T.R.
Proc. Natl. Acad. Sci. U.S.A. 95, 8479-8484, 1998
A:Title: Telomerase reverse transcriptase genes in Tetrahymena thermophila and Oxytricha
A:Reference number: Z20985; MUID:98337940
A:Accession: T31107
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1132 <BRY>
A:Cross-references: EMBL:AF060230; NID:g3342795; PID:g3342796; PIDN:AAC39163.1
C:Genetics:
A:Gene: TERT

Query Match 80.0%; Score 36; DB 2; Length 1132;
Best Local Similarity 75.0%; Pred. No. 50;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NNATFYFK 8
| :|||||
Db 1001 NNISFYFK 1008

RESULT 7
T24206
hypothetical protein R12H7.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 08-Dec-2000
C:Accession: T24206
R:Coles, L.
submitted to the EMBL Data Library, August 1995
A:Reference number: Z19854
A:Accession: T24206
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-182 <WIL>
A:Cross-references: EMBL:Z50755; PIDN:CAA90635.1; GSPDB:GN00028; CESP:R12H7.3
A:Experimental source: clone R12H7
C:Genetics:
A:Gene: CESP:R12H7.3
A:Map position: X
A:Introns: 150/3

C:Superfamily: human S-phase kinase-associated protein 1A

Query Match 77.8%; Score 35; DB 2; Length 182;
Best Local Similarity 75.0%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NNATFYFK 8
| :|||||
Db 161 NNATLFFK 168

RESULT 8
B83803
hypothetical protein BHI226 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: B83803
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: B83803
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-71 <STO>
A:Cross-references: GB:AP001511; GB:BA000004; NID:g10173727; PIDN:BA04945.1; GSPDB:G
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BHI226

Query Match 75.6%; Score 34; DB 2; Length 71;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NATFYF 7
| :|||||
Db 43 NATFYF 48

RESULT 9
T06978
ABA-induced plasma membrane protein PM 19 - wheat
C:Species: Triticum aestivum (common wheat)
C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 08-Oct-1999
C:Accession: T06978
R:Koike, M.; Takezawa, D.; Arakawa, K.; Yoshida, S.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z15842
A:Accession: T06978
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-182 <KOI>
A:Cross-references: EMBL:U80037; NID:g1724111; PIDN:AAB38504.1; PID:g1724112
A:Experimental source: cv. Chihoku
C:Genetics:
A:Note: WTABAPM

Query Match 75.6%; Score 34; DB 2; Length 182;
Best Local Similarity 85.7%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NNATFYF 7
| :|||||
Db 47 NGATFYF 53

RESULT 10
F86924
hypothetical protein [imported] - Mycobacterium leprae
C:Species: Mycobacterium leprae

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C;Accession: F86924
R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Holt,
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Squ
A;Title: Massive gene decay in the leprosy bacillus.
A;Reference number: A86909; MUID:21128732; PMID:11234002
A;Accession: F86924
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-273 <STO>
A;Cross-references: GB:AL450380; NID:gl3092504; PIDN:CAC29634.1; GSPDB:GN00147
C;Genetics:
A;Gene: ML0126

```

Query Match          75.6%; Score 34; DB 2; Length 273;
Best Local Similarity 85.7%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNATFYF 7
    ||| |||
Db 20 NNAQFYF 26

RESULT 11
A86650
Rhamnosyltransferase [Imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: A86650
R:Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Mialme, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: A86650
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1319 <STO>
A:Cross-references: GB:AE005176; PID:g12723056; PIDN:AAK04299.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: rqbB

```

```

Query Match          75.6%; Score 34; DB 2; Length 319;
Best Local Similarity 85.7%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 NNATYF 7
        ||| |||
Db       80 NNADYF 86

RESULT 12
T11319
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 2 - Pedinomonas minor mitochondrion
C:Species: mitochondrion Pedinomonas minor
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 17-Mar-2000
C:Accession: T11319
R:Turmel, M.; Lemieux, C.; Burger, G.; Lang, B.F.; Otis, C.; Plante, I.; Gray, M.W.
submitted to the EMBL Data Library, December 1998
A:Description: The complete mitochondrial DNA sequences of Nephroselmis olivacea and Ped
A:Reference number: Z17261
A:Accession: T11319
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-440 <TUR>
A:Cross-references: EMBL:AF116775; NID:g4378766; PID:g4378773; PIDN:AAD19671.1
C:Genetics:
A:Gene: nad2

```

A: Genome: mitochondrion
A: Genetic code: SCC3
C: Superfamily: NADH dehydrogenase (ubiquinone) chain 2
C: Keywords: mitochondrion; NAD; oxidoreductase

Query Match 75.6%; Score 34; DB 2; Length 440;
Best Local Similarity 85.7%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NNATFYF 7
 | | | | |
Db 238 NNALFYF 244

RESULT 13
G96611
Probable cytochrome P450 T9L23.21 [imported] - Arabidopsis thaliana
C: Species: Arabidopsis thaliana (mouse-ear cress)
C: Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C: Accession: G96611
R: Theologis, A.; Ecker, J. R.; Palm, C. J.; Federspiel, N. A.; Kaul, S.; White, O.; Alon
Chn, C. W.; Chung, M. K.; Conn, L.; Conway, A. B.; Conway, A. R.; Creasy, T. H.; Dewar,
ansen, N. F.; Hughes, B.; Huizlar, L.
Nature 408, 816-820, 2000
A: Authors: Hunter, J. L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C. A.; Li, J. H.; Li, Y.; Lin, X.; Liu, Z. A.; Luros, J. S.; Maiti, R.; Marzita
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A: Authors: Salzberg, S. L.; Schwartz, J. R.; Shinn, P.; Southwick, A. M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C. M.; Venter, J. C.; Davis, R. W.
A: Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A: Reference number: A86141; MUID: 21016719

```

Query Match          75.6%; Score 34; DB 2; Length 497;
Best Local Similarity 85.7%; Pred. No. 54;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 NATYFK 8
      | | | | |
Db      64 NLTFYFK 70

RESULT 14
T25830
hypothetical protein M01A10.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T25830
R:Sheet, P.
submitted to the EMBL Data Library, February 1997
A:Description: The sequence of C. elegans cosmid M01A10.
A:Reference number: Z20094
A:Accession: T25830
A:Status: preliminary; translated from GB/EMBL/DBRJ
A:Molecule type: DNA
A:Residues: 1-630 <SCH>
A:Cross-references: EMBL:U88174; PIDN:AAB42276.1; GSPDB:GN00019; CESP:M01A10.4
A:Experimental source: strain Bristol N2; clone M01A10
C:Genetics:
A:Gene: CESP:M01A10.4
A:Map position: 1
A:Introns: 10/2; 249/3; 284/1

```

Query Match 75.6%; Score 34; DB 2; Length 630;
Best Local Similarity 75.0%; Pred. No. 68;
Matches 6; Conservative 1; Mismatches 0; Gaps 0;

Qy 1 NNATFYFK 8
: || ||||
Db 444 SNAQFYFK 451

RESULT 15

T31994
hypothetical protein C49D10.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T31994
R:Henkhaus, J.; Wohldmann, P.; Beck, C.
submitted to the EMBL Data Library, July 1997
A:Description: The sequence of C. elegans cosmid C49D10.
A:Reference number: Z21108
A:Accession: T31994
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-758 <HEN>
A:Cross-references: EMBL:AF016665; PIDN:AACT1186.1; GSPDB:GN00020; CESP:C49D10.1
A:Experimental source: strain Bristol N2; clone C49D10
C:Genetics:
A:Gene: CESP:C49D10.1
A:Map position: 2
A:Introns: 438/2

Query Match 75.6%; Score 34; DB 2; Length 758;
Best Local Similarity 75.0%; Pred. No. 82;
Matches 6; Conservative 1; Mismatches 0; Gaps 0;

Qy 1 NNATFYFK 8
: || ||||
Db 598 SNAQFYFK 605

Search completed: July 10, 2002, 09:21:18
Job time: 44 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2002, 09:20:34 ; Search time 10.34 Seconds
(without alignments)
29.957 Million cell updates/sec

Title: US-09-461-061A-1
Perfect score: 45
Sequence: 1 NNATFFYK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	45	100.0	644	1 KNG_HUMAN	P01042 homo sapien
2	39	86.7	639	1 KNG_RAT	P08934 rattus norv
3	36	80.0	1132	1 TERT_OXYTR	Q07332 oxytricha t
4	33	73.3	603	1 CHLE_MOUSE	Q03311 mus musculu
5	32	71.1	562	1 INVG_SALTY	P35642 salmonella
6	32	71.1	1590	1 GCM2_YEAST	P15442 saccharomyc
7	31	68.9	82	1 U2AG_PIG	Q29350 sus scrofa
8	31	68.9	196	1 HIS5_BUCAI	P57204 buchnera ap
9	31	68.9	239	1 U2AG_MOUSE	Q9d883 mus musculu
10	31	68.9	240	1 U2AG_HUMAN	Q01081 homo sapien
11	31	68.9	264	1 U2AG_DROME	Q94535 drosophila
12	31	68.9	334	1 ARGC_ECOLI	P11446 escherichia
13	31	68.9	341	1 OMPU_VIBCH	P97085 vibrio chol
14	31	68.9	434	1 KNL2_BOVIN	P01047 bos taurus
15	31	68.9	441	1 PAFA_HUMAN	Q13093 h platelet-
16	31	68.9	444	1 PAFA_CANFA	Q28262 c platelet-
17	31	68.9	619	1 KNH2_BOVIN	P01045 bos taurus
18	31	68.9	623	1 RSD1_YEAST	P32368 saccharomyc
19	31	68.9	901	1 VEF_GVPU	P41723 pseudolatia
20	31	68.9	901	1 VEF_GVTN	P29998 trichoplusi
21	31	68.9	902	1 VEF_GVHA	P54232 heliothis a
22	31	68.9	988	1 ST23_YEAST	Q06010 saccharomyc
23	31	68.9	1066	1 NUC2_NEUCR	Q01317 neurospora
24	31	68.9	2710	1 TOXA_CLODI	P16154 clostidium
25	31	68.9	4725	1 DYHC_DICDI	P34036 dictyostell
26	30	66.7	163	1 SFAS_ECOLI	P13430 escherichia
27	30	66.7	196	1 HIS5_METJA	Q57929 methanococc
28	30	66.7	208	1 LEUD_BUCDN	O85073 buchnera ap
29	30	66.7	237	1 YS76_CAEEL	Q09619 caenorhabdi
30	30	66.7	259	1 KKA6_ACTIB	P09885 acinetobact
31	30	66.7	331	1 NIXA_HELPJ	Q9zm74 helicobacte
32	30	66.7	331	1 NIXA_HELPY	Q48262 helicobacte
33	30	66.7	365	1 NMPC_ECOLI	P21420 escherichia

34	30	66.7	383	1 CYB_APLLI	P34845 apis mellif
35	30	66.7	397	1 YAK7_SCHPO	Q09919 schizosacch
36	30	66.7	452	1 SMP1_YEAST	P38128 saccharomyc
37	30	66.7	477	1 GLGA_STRPN	O97855 streptococc
38	30	66.7	492	1 CPBU_MOUSE	O5071 mus musculu
39	30	66.7	537	1 YIV9_YEAST	P40383 saccharomyc
40	30	66.7	541	1 UL21_VZVD	P09289 varicella-z
41	30	66.7	581	1 CHLE_RABIT	P21927 oryctolagus
42	30	66.7	602	1 CHLE_HUMAN	P06276 homo sapien
43	30	66.7	603	1 NUSM_MYOGL	O63908 myoxus glis
44	30	66.7	758	1 LZDT_DROME	Q24371 drosophila
45	30	66.7	952	1 PM16_CHLPN	Q92882 chlamydia p

ALIGNMENTS

RESULT 1
KNG_HUMAN
ID KNG_HUMAN STANDARD; PRT; 644 AA.
AC P01042: P01043;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Kininogen precursor (Alpha-2-thiol proteinase inhibitor) [Contains:
DE Bradykinin].
GN KNG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
RC TISSUE=Liver; PubMed=2989293;
RX MEDLINE=85234582; PubMed=2989293;
RA Takagaki Y., Kitamura N., Nakanishi S.;
RT "Cloning and sequence analysis of cDNAs for human high molecular
RT weight and low molecular weight prekininogens. Primary structures of
RT two human prekininogens".
RL J. Biol. Chem. 260:8601-8609(1985).
RN [2]
RP GENE STRUCTURE.
RX MEDLINE=85234583; PubMed=2989294;
RA Kitamura N., Kitagawa H., Fukushima D., Takagaki Y., Miyata T.,
RA Nakanishi S.;
RT "Structural organization of the human kininogen gene and a model for
RT its evolution".
RL J. Biol. Chem. 260:8610-8617(1985).
RN [3]
RP SEQUENCE OF 1-401 FROM N.A.
RX MEDLINE=85122621; PubMed=6441591;
RA Ohkubo I., Kurachi K., Takasawa T., Shiokawa H., Sasaki M.;
RT "Isolation of a human cDNA for alpha 2-thiol proteinase inhibitor and
RL its identity with low molecular weight kininogen".
RN [4]
RP Biochemistry 23:5691-5697(1984).
RX SEQUENCE OF 379-644.
RP MEDLINE=86030270; PubMed=4054110;
RA Lottspeich F., Kellermann J., Henschel A., Foertsch B.,
RA Mueller-Esterl W.;
RT "The amino acid sequence of the light chain of human high-molecular-
RT mass kininogen".
RL Eur. J. Biochem. 152:307-314(1985).
RN [5]
RP SEQUENCE OF 381-389.
RX MEDLINE=90255622; PubMed=4952632;
RA Pierce J.V.;
RT "Structural features of plasma kinins and kininogens".
RL Red. Proc. 27:52-57(1968).
RN [6]
RP DISULFIDE BONDS.
RA Sueyoshi T., Miyata T., Kato H., Iwanaga S.;
RT "Disulfide bonds in bovine HMW kininogens.";

RL Seikagaku 56:808-808(1984).
 CC -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
 CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
 CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO
 CC FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN-AND PLASMIN-
 CC INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE
 CC PHYSIOLOGICAL THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF
 CC PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE
 CC CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NATRIURESIS AND
 CC DIURESIS, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A
 CC MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR
 CC PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF
 CC OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS
 CC A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION, (5)
 CC INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)
 CC LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-
 CC KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD
 CC CLOTTING.
 CC
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- TISSUE SPECIFICITY: PLASMA.
 CC -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
 CC -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; K02566; AAA35497.1; -
 CC EMBL; M11437; AAB59550.1; -
 CC EMBL; M11438; AAB59550.1; JOINED.
 CC EMBL; M11521; AAB59550.1; JOINED.
 CC EMBL; M11522; AAB59550.1; JOINED.
 CC EMBL; M11523; AAB59550.1; JOINED.
 CC EMBL; M11524; AAB59550.1; JOINED.
 CC EMBL; M11525; AAB59550.1; JOINED.
 CC EMBL; M11526; AAB59550.1; JOINED.
 CC EMBL; M11527; AAB59550.1; JOINED.
 CC EMBL; M11528; AAB59550.1; JOINED.
 CC EMBL; M11437; AAB59551.1; -
 CC EMBL; M11438; AAB59551.1; JOINED.
 CC EMBL; M11521; AAB59551.1; JOINED.
 CC EMBL; M11522; AAB59551.1; JOINED.
 CC EMBL; M11523; AAB59551.1; JOINED.
 CC EMBL; M11524; AAB59551.1; JOINED.
 CC EMBL; M11525; AAB59551.1; JOINED.
 CC EMBL; M11526; AAB59551.1; JOINED.
 CC EMBL; M11527; AAB59551.1; JOINED.
 CC EMBL; M11528; AAB59551.1; JOINED.
 CC PIR; A01279; KGRUHL.
 CC PIR; A25276; A25276.
 CC PIR; A01280; KGHUL1.
 CC PIR; B25276; B25276.
 CC PIR; S02482; S02482.
 CC SWISS-2DPAGE; P01042; HUMAN.
 CC MIM; 228960; -
 CC InterPro; IPR000010; Cystatin.
 CC InterPro; IPR003243; Cystatin_C.M.
 CC InterPro; IPR002395; Kininogen.
 CC Pfam; PF00031; cystatin; 3.
 CC PRINTS; PR00334; KININOGEN.
 CC Prodom; PD001231; Cystatin_C.M; 1.
 CC SMART; SM00043; CY; 3.
 CC PROSITE; PS00287; CYSTATIN; 2.
 CC Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
 CC Bradykinin; Blood coagulation; Inflammatory response; Signal;
 CC Alternative splicing.
 CC SIGNAL 1 18

FT CHAIN 19 644 KININOGEN.
 FT CHAIN 19 380 KININOGEN HEAVY CHAIN.
 FT PEPTIDE 381 389 BRADYKININ.
 FT CHAIN 390 644 KININOGEN LIGHT CHAIN.
 FT DOMAIN 19 136 CYSTATIN-LIKE 1.
 FT DOMAIN 137 258 CYSTATIN-LIKE 2.
 FT DOMAIN 259 380 CYSTATIN-LIKE 3.
 FT DOMAIN 420 510 HIS-RICH (ASSOCIATED WITH CLOTTING
 FT REPEAT 420 449 ACTIVITY).
 FT REPEAT 450 479
 FT REPEAT 480 510
 FT MOD_RES 19 19 PYRROLIDONE CARBOXYLIC ACID.
 FT DISULFID 28 614 INTERCHAIN.
 FT DISULFID 83 94
 FT DISULFID 107 126
 FT DISULFID 142 145
 FT DISULFID 206 218
 FT DISULFID 229 248
 FT DISULFID 264 267
 FT DISULFID 328 340
 FT DISULFID 351 370
 FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 401 401 O-LINKED.
 FT CARBOHYD 533 533 O-LINKED.
 FT CARBOHYD 542 542 O-LINKED.
 FT CARBOHYD 546 546 O-LINKED.
 FT CARBOHYD 557 557 O-LINKED.
 FT CARBOHYD 571 571 O-LINKED.
 FT CARBOHYD 577 577 O-LINKED.
 FT CARBOHYD 593 593 O-LINKED.
 FT CARBOHYD 628 628 O-LINKED.
 FT VARSPLIC 402 427
 FT VARSPLIC 428 644
 FT CONFLICT 593 593 MISSING (IN ISOFORM LMW).
 FT CONFLICT 593 593 T -> I (IN REF. 1).
 SQ SEQUENCE 644 AA; 71945 MW; 3132B4CBFAF8FB7E CRC64;
 Query Match 100.0%; Score 45; DB 1; Length 644;
 Best Local Similarity 100.0%; Pred. No. 0.24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NNATYFK 8
 Db 293 NNATYFK 300
 RESULT 2
 KNG_RAT
 ID KNG_RAT STANDARD; PRT; 639 AA.
 AC P08934; P08933;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Kininogen precursor [Contains: Bradykinin].
 GN KNG.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS HMW AND LMW).
 RX MEDLINE=87137443; PubMed=3029068;
 RA Kitagawa H., Kitamura N., Hayashida H., Miyata T., Nakanishi S.;
 RT "Differing expression patterns and evolution of the rat kininogen
 RT gene family";
 RL J. Biol. Chem. 262:2190-2198(1987).
 RN [2]
 RP SEQUENCE FROM N.A. (LMW ISOFORM).

RA MEDLINE=86008264; PubMed=2413018;
RA Furuto-Kato S., Matsumoto A., Kitamura N., Nakanishi S.;
RT "Primary structures of the mRNAs encoding the rat precursors for
RT bradykinin and T-kinin. Structural relationship of kininogens with
RT major acute phase protein and alpha 1-cysteine proteinase
RT inhibitor.";
RL J. Biol. Chem. 260:12054-12059(1985).
RN [3]
RP SEQUENCE OF 1-65 FROM N.A.
RC STRAIN=BUFEALO;
RX MEDLINE=87250580; PubMed=2439509;
RA Fung W.-P., Schreiber G.;
RT "Structure and expression of the genes for major acute phase alpha 1-
RT protein (thiostatin) and kininogen in the rat.";
RL J. Biol. Chem. 262:9298-9308(1987).
RN [4]
RP SEQUENCE OF 1-41 FROM N.A.
RC STRAIN=WISTAR; TISSUE=Liver;
RX MEDLINE=87137465; PubMed=3818598;
RA Kageyama R., Kitamura N., Ohkubo H., Nakanishi S.;
RT "Differing utilization of homologous transcription initiation sites
RT of rat K and T kininogen genes under inflammation condition.";
RL J. Biol. Chem. 262:2345-2351(1987).
CC -!- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC HMW-KININOGEN PLAYS AN IMPORTANT ROLE IN BLOOD COAGULATION BY
CC HELPING TO POSITION OPTIMALLY PREKALLIKREIN AND FACTOR XI NEXT TO
CC FACTOR XII; (3) HMW-KININOGEN INHIBITS THE THROMBIN-AND PLASMIN-
CC INDUCED AGGREGATION OF THROMBOCYTES; (4) THE ACTIVE PEPTIDE
CC BRADYKININ THAT IS RELEASED FROM HMW-KININOGEN SHOWS A VARIETY OF
CC PHYSIOLOGICAL EFFECTS: (4A) INFLUENCE IN SMOOTH MUSCLE
CC CONTRACTION, (4B) INDUCTION OF HYPOTENSION, (4C) NUTRIURETIC AND
CC DIURETIC, (4D) DECREASE IN BLOOD GLUCOSE LEVEL, (4E) IT IS A
CC MEDIATOR OF INFLAMMATION AND CAUSES (4E1) INCREASE IN VASCULAR
CC PERMEABILITY, (4E2) STIMULATION OF NOCICEPTORS (4E3) RELEASE OF
CC OTHER MEDIATORS OF INFLAMMATION (E.G. PROSTAGLANDINS), (4F) IT HAS
CC A CARDIOPROTECTIVE EFFECT (DIRECTLY VIA BRADYKININ ACTION,
CC INDIRECTLY VIA ENDOTHELIUM-DERIVED RELAXING FACTOR ACTION); (5)
CC HMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (6) LMW-
CC KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT INVOLVED IN BLOOD
CC CLOTTING.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; HMW (SHOWN HERE) AND LMW; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: PLASMA.
CC -!- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC -!- MISCELLANEOUS: RAT EXPRESS FOUR TYPES OF KININOGENS: THE CLASSICAL
CC HMW/LMW KININOGENS AND TWO ADDITIONAL LMW-LIKE KININOGENS: T-I AND
CC T-II.
CC -!- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L29428; AAA41486.1; -;
DR EMBL; M11884; AAA41487.1; -;
DR EMBL; M14369; AAA41484.1; -;
DR EMBL; M14369; AAA41485.1; ALT_SEQ.
DR EMBL; M16455; AAA41482.1; -;
DR PIR; A25486; A25486.
DR PIR; A28055; A28055.
DR HSSP; P01040; 1DND.
DR InterPro; IPR000010; Cystatin.
DR InterPro; IPR003243; Cystatin_C_M.
DR InterPro; IPR002395; Kininogen.
DR Pfam; PF00031; cystatin; 3.
DR PRINTS; PR00334; KININOGEN.
DR ProDom; PD001231; Kininogen.
DR SMART; SM00043; CY; 3.

DR PROSITE; PS00287; CYSTATIN; 2.
KW Glycoprotein; Plasma; Repeat; Thiol protease inhibitor; Vasodilator;
KW Bradykinin; Blood coagulation; Inflammatory response; Signal;
KW Alternative splicing; Multigene family.
FT SIGNAL 1 18
FT CHAIN 19 639 KININOGEN.
FT CHAIN 19 380 KININOGEN HEAVY CHAIN.
FT PEPTIDE 381 389 BRADYKININ.
FT CHAIN 390 639 KININOGEN LIGHT CHAIN.
FT DOMAIN 19 136 CYSTATIN-LIKE 1.
FT DOMAIN 137 258 CYSTATIN-LIKE 2.
FT DOMAIN 259 380 CYSTATIN-LIKE 3.
FT DOMAIN 439 514 HIS-RICH.
FT DISULFID 28 609 INTERCHAIN (BY SIMILARITY).
FT DISULFID 83 94 BY SIMILARITY.
FT DISULFID 107 126 BY SIMILARITY.
FT DISULFID 142 145 BY SIMILARITY.
FT DISULFID 206 218 BY SIMILARITY.
FT DISULFID 229 248 BY SIMILARITY.
FT DISULFID 264 267 BY SIMILARITY.
FT DISULFID 328 340 BY SIMILARITY.
FT DISULFID 351 370 BY SIMILARITY.
FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 529 529 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 402 433 VSPSIARVQERDPEGNGPIHGHWLHAKQ -> RLNS
FT CEYKGRLLKAGAPAPQEAESTVTP (IN ISOFORM
FT LMW).
FT VARSPLIC 434 639 MISSING (IN ISOFORM LMW).
FT CONFLICT 61 61 E -> K (IN REF. 2).
FT SEQUENCE 639 AA; 70933 MW; D3172DF94FF56AF5 CRC64;
SQ
Query Match 86.7%; Score 39; DB 1; Length 639;
Best Local Similarity 87.5%; Pred. No. 3.3;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 NNATFYFK 8
DB 293 NNHTFYFK 300
RESULT 3
TERT_OXYTR
ID TERT_OXYTR STANDARD; PRT: 1132 AA.
AC O76332;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Telomerase reverse transcriptase (EC 2.7.7.-) (Telomerase catalytic
GN subunit) (Telomerase subunit P133).
OS Oxytricha trifallax.
OC Eukaryota; Alveolata; Ciliophora; hypotrichs; Stichotrichida;
OC Oxytrichidae; Oxytricha.
OX NCBI_Taxid=5946;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98337940; PubMed=9671703;
RA Bryan T.M., Sperger J.M., Chapman K.B., Cech T.R.;
RT "Telomerase reverse transcriptase genes identified in Tetrahymena
RT thermophila and Oxytricha trifallax.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:8479-8484(1998).
CC -!- FUNCTION: TELOMERASE IS A RIBONUCLEOPROTEIN ENZYME ESSENTIAL FOR
CC THE REPLICATION OF CHROMOSOME TERMINI IN MOST EUKARYOTES. IT
CC ELONGATES TELOMERES. IT IS A REVERSE TRANSCRIPTASE THAT ADDS
CC SIMPLE SEQUENCE REPEATS TO CHROMOSOME ENDS BY COPYING A TEMPLATE
CC SEQUENCE WITHIN THE RNA COMPONENT OF THE ENZYME.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE REVERSE TRANSCRIPTASE FAMILY.

TELOMERASE SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; AF060230; AAC39163.1; -
 InterPro; IPR003545; Telomerase_RT.
 PRINTS; PR01365; TELOMERASE_RT.
 Transferase; RNA-directed DNA polymerase; Telomere; Nuclear protein;
 KW DNA-binding
 QY SEQUENCE 1132 AA; 134124 MW; 81E145F5F24392DC CRC64;

Query Match 80.0%; Score 36; DB 1; Length 1132;
 Best Local Similarity 75.0%; Pred. No. 22;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NNATFYFK 8
 || : ||||
 Db 1001 NNISFYFK 1008

RESULT 4
 ID CHLE_MOUSE STANDARD; PRT; 603 AA.
 AC Q03311;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Cholinesterase precursor (EC 3.1.1.8) (Acylocholine acylhydrolase)
 DE (Choline esterase II) (Butyrylcholine esterase)
 DE (Pseudochoolinesterase).
 GN BCHE.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90380429; PubMed=2400605;
 RA Rachinsky T.L., Camp S., Li Y., Ekstrom T.J., Newton M., Taylor P.;
 RT "Molecular cloning of mouse acetylcholinesterase: tissue distribution of alternatively spliced mRNA species.";
 RL Neuron 5:317-327(1990).
 RN [2]
 RP SEQUENCE OF 97-237 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=91201348; PubMed=2016308;
 RA Arpagaus M., Chatonnet A., Masson P., Newton M., Vaughan T.A.,
 RA Bartels C.F., Nogueira C.P., la Du B.N., Lockridge O.;
 RT "Use of the polymerase chain reaction for homology probing of butyrylcholinesterase from several vertebrates.";
 RL J. Biol. Chem. 266:6966-6974(1991).
 CC -I- CATALYTIC ACTIVITY: An acylcholine + H(2)O = choline + a carboxylic acid anion.
 CC -I- SUBUNIT: HOMOTETRAMER. THE TETRAMER IS COMPOSED OF TWO DIMERS. THE TWO SUBUNITS IN A DIMER ARE LINKED BY A DISULFIDE BOND.
 CC -I- TISSUE SPECIFICITY: PRESENT IN MOST CELLS (EXCEPT ERYTHROCYTES).
 CC -I- MISCELLANEOUS: CHOLINESTERASE IS HIGHLY REACTIVE WITH ORGANOPHOSPHATE ESTERS.
 CC -I- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; M99492; AAA37328.1; -
 PIR; A39768; A39768.
 DR HSP; P21836; IMAH.
 DR MGD; MGI:894278; Bche.
 DR InterPro; IPR002018; Carboxylesterase_B.
 DR InterPro; IPR000997; Cholinesterase.
 DR InterPro; IPR000379; Est_lip_thioest_actsite.
 DR Pfam; PF00135; Coesterase; 1.
 DR PRINTS; PR00878; CHOLINESTRASE.
 DR PROSITE; PS00122; CARBOXYLESTERASE_B.1; 1.
 DR PROSITE; PS00941; CARBOXYLESTERASE_B.2; 1.
 KW Hydrolase; Serine esterase; Glycoprotein; Signal.
 FT SIGNAL 1 29
 FT CHAIN 30 603 CHOLINESTRASE.
 FT ACT_SITE 227 227 BY SIMILARITY.
 FT ACT_SITE 354 354 BY SIMILARITY.
 FT ACT_SITE 467 467 BY SIMILARITY.
 FT DISULFID 94 121 BY SIMILARITY.
 FT DISULFID 281 292 BY SIMILARITY.
 FT DISULFID 429 548 BY SIMILARITY.
 FT DISULFID 600 600 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 86 86 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 270 270 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 370 370 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 484 484 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 510 510 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 515 515 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 129 129 R -> P (IN REF. 2).
 SQ SEQUENCE 603 AA; 68521 MW; 719B1B220D1E5367 CRC64;

Query Match 73.3%; Score 33; DB 1; Length 603;
 Best Local Similarity 85.7%; Pred. No. 44;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NNATFYF 7
 ||| |||
 Db 443 NNATFYF 449

RESULT 5
 ID INVG_SALTY STANDARD; PRT; 562 AA.
 AC P35672;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE INVG protein precursor.
 GN INVG OR STM2898.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TWL;
 RX MEDLINE=95117219; PubMed=7868245;
 RA Lodge J., Douce G.R., Amin I.I., Bolton A.J., Martin G.D.,
 RA Chatfield S., Dougan G., Brown N.L., Stephen J.;
 RT "Biological and genetic characterization of TphoA mutants of Salmonella typhimurium TML in the context of gastroenteritis.";
 RL Infect. Immun. 63:762-769(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SR-11;
 RX MEDLINE=95089692; PubMed=7997169;
 RA Kaniga K., Bossio J.C., Galan J.E.;
 RT "The Salmonella typhimurium invasion genes invF and invG encode homologues of the AraC and PvdL family of proteins.";

RL Mol. Microbiol. 13:555-568(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE-21534948; PubMed-11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 LT2.";
 RL Nature 413:852-856(2001).
 RC STRAIN-S288C;
 CC -!- FUNCTION: INVOLVED IN THE INVASION OF THE CELLS OF THE INTESTINAL
 CC EPITHELIUM. COULD BE NECESSARY FOR THE EXPORT OF INVASION RELATED
 CC DETERMINANTS.
 CC -!- SUBCELLULAR LOCATION: Outer membrane (Potential).
 CC -!- SIMILARITY: BELONGS TO THE PULD/OUTD/EXED/XPSD FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X75302; CAAS3049.1; -;
 DR EMBL; U08280; AAA74040.1; -;
 DR EMBL; AE008832; AAL21778.1; -;
 DR PIR; S38422; S38422.
 DR StyGene; SG10188; invG.
 DR InterPro; IPR000016; Bac-GSPproteins.
 DR Pfam; PF00263; GSP11_III; 1.
 DR PRINTS; PR01337; TYPE3OMGPROT.
 DR PROSITE; PS00875; T2SP_D; 1.
 DR Virulence; Transports; Protein transport; Signal; Outer membrane;
 KW Complete proteome.
 FT SIGNAL 1 14 POTENTIAL.
 FT CHAIN 15 562 INV G PROTEIN.
 FT CONFLICT 12 12 A -> R (IN REF. 2).
 FT CONFLICT 121 121 E -> Q (IN REF. 2).
 FT CONFLICT 197 205 LRQKQVVP -> CAIRKWLFR (IN REF. 2).
 FT CONFLICT 232 240 AMPAFSANG -> RCQRFQM (IN REF. 2).
 FT CONFLICT 243 243 G -> S (IN REF. 2).
 FT CONFLICT 262 264 AAA -> KPAEQ (IN REF. 2).
 FT CONFLICT 328 328 S -> T (IN REF. 1 AND 2).
 FT CONFLICT 329 329 I -> V (IN REF. 2).
 FT CONFLICT 370 380 RPLVLTQENVP -> APTGTSKGCS (IN REF. 2).
 SQ SEQUENCE 562 AA; 61765 MW; 80222905BE256058D CRC64;

 Query Match 71.1%; Score 32; DB 1; Length 562;
 Best Local Similarity 75.0%; Pred. No. 63;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

 QY 1 NNATFYFK 8
 II III I
 DB 385 NNRTFYTK 392

 RESULT 6
 GCN2_YEAST
 ID GCN2_YEAST STANDARD; PRT; 1590 AA.
 AC P15442;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Protein kinase GCN2 (EC 2.7.1.-).
 GN GCN2 OR AAS1 OR YDR283C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_taxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-89282814; PubMed-2660141;
 RX WEK R.C., JACKSON B.M., HINNEBUSCH A.G.;
 RA "Juxtaposition of domains homologous to protein kinases and histidyl-
 RT tRNA synthetases in GCN2 protein suggests a mechanism for coupling
 RT GCN4 expression to amino acid availability.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:4579-4583(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C;
 RX MEDLINE-89282814; PubMed-3290651;
 RA ROUSSOU I., THIROOS G., HAUGE B.M.;
 RT "Transcriptional-translational regulatory circuit in Saccharomyces
 RT cerevisiae which involves the GCN4 transcriptional activator and the
 RT GCN2 protein kinase.";
 RL Mol. Cell. Biol. 8:2132-2139(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C / AB972;
 RA Le T.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP ASSOCIATION WITH RIBOSOMES.
 RX MEDLINE-91246169; PubMed-2038314;
 RA Ramirez M., WEK R.C., HINNEBUSCH A.G.;
 RT "Ribosome association of GCN2 protein kinase, a translational
 RT activator of the GCN4 gene of Saccharomyces cerevisiae.";
 RL Mol. Cell. Biol. 11:3027-3036(1991).
 RN [5]
 RP ACTIVITY ON SUI2.
 RX MEDLINE-92154672; PubMed-1739968;
 RA Dever T.E., Feng L., WEK R.C., CIGAN A.M., DONAHUE T.D.,
 RA HINNEBUSCH A.G.;
 RT "Phosphorylation of initiation factor 2 alpha by protein kinase GCN2
 RT mediates gene-specific translational control of GCN4 in yeast.";
 RL Cell 68:585-596(1992).
 CC -!- FUNCTION: STIMULATES GCN4 TRANSLATION IN AMINO ACID-STARVED CELLS
 CC BY PHOSPHORYLATING THE ALPHA SUBUNIT OF EIF-2 (SUI2) ON SER-52.
 CC -!- SUBCELLULAR LOCATION: ASSOCIATES WITH THE 60S SUBUNIT OF
 CC RIBOSOMES.
 CC -!- INDUCTION: BY AMINO ACID STARVATION AND IN THE PRESENCE OF THE
 CC GCN4 PROTEIN.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC GCN2 SUBFAMILY.
 CC -!- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN IN POSITIONS 301
 CC TO 364 AND FROM POSITION 981 ONWARD DUE TO FRAMESHIFTS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M27082; AAA34636.1; -;
 DR EMBL; M20487; AAA34881.1; ALT_FRAME.
 DR EMBL; U51030; AAB64461.1; ALT_INIT.
 DR PIR; S05781; OKBYN2.
 DR PIR; A27723; A27723.
 DR HSP; P24941; 1JST.
 DR SGD; S0002691; GCN2.
 DR InterPro; IPR002106; AA_trna_ligase_II.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR Pfam; PF00069; pkinase; 4.
 DR PROSITE; PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 2.

KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Amino-acid biosynthesis; Aminoacyl-tRNA synthetase; Repeat.
FT DOMAIN 187 458 PROTEIN KINASE 1.
FT DOMAIN 530 912 PROTEIN KINASE 2.
FT DOMAIN 930 1450 HISTIDYL-TRNA SYNTHETASE LIKE.
FT NP_BIND 536 544 ATP (BY SIMILARITY).
FT BINDING 559 559 ATP (BY SIMILARITY).
FT ACT_SITE 766 766 TIRAKLP -> NYKQKIA (IN REF. 2).
FT CONFLICT 151 157 LMSSEMEN -> YVFSNHGKS (IN REF. 2).
FT CONFLICT 202 210 I -> M (IN REF. 2).
FT CONFLICT 406 406 P -> A (IN REF. 2).
FT CONFLICT 520 520 RSYASDFEETAVL -> DHDMLTLTKRLHFS
FT CONFLICT 523 536 (IN REF. 2).
FT CONFLICT 553 553 S -> T (IN REF. 2).
FT CONFLICT 571 572 IL -> MI (IN REF. 2).
FT CONFLICT 658 658 F -> C (IN REF. 2).
FT CONFLICT 770 770 M -> K (IN REF. 2).
FT CONFLICT 885 885 E -> Q (IN REF. 2).
SQ SEQUENCE 1590 AA; 182041 MW; 768815424F64F1AA CRC64;

Query Match 71.18; Score 32; DB 1; Length 1590;
Best Local Similarity 75.08; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NNATFYFK 8

|||||:

Db 266 NNATFVWK 273

RESULT 7

ID U2AG_PIG STANDARD; PRT; 82 AA.
AC Q29350;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Splicing factor U2AF 35 kDa subunit (U2 auxiliary factor 35 kDa
DE subunit) (U2 snRNP auxiliary factor small subunit) (Fragment).
GN U2AF1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RX MEDLINE=96327607; PubMed=8672129;
RA Winteroe A.K., Fredholm M., Davies W.;
RT "Evaluation and characterization of a porcine small intestine cDNA
RT library: analysis of 839 clones.";
RL Mamm. Genome 7:509-517(1996).
CC -1- FUNCTION: PLAYS A CRITICAL ROLE IN BOTH CONSTITUTIVE AND ENHANCER-
CC DEPENDENT SPLICING BY MEDIATING ESSENTIAL PROTEIN-PROTEIN
CC INTERACTIONS AND PROTEIN-RNA INTERACTIONS REQUIRED FOR ACCURATE 3'
CC SPLICE SITE SELECTION. DIRECTLY MEDIATES INTERACTIONS BETWEEN
CC U2AF65 (LARGE SUBUNIT) AND PROTEINS BOUND TO THE ENHANCERS AND
CC THIS MAY FUNCTION AS A BRIDGE BETWEEN U2AF65 AND THE ENHANCER
CC COMPLEX TO RECRUIT IT TO THE ADJACENT INTRON (BY SIMILARITY).
CC -1- SUBUNIT: ASSOCIATES WITH A 65 KDA PROTEIN.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -1- SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

DR EMBL; F14757; CAA23231.1; -.
DR InterPro; IPR000571; zf-CCCH.
DR Pfam; PF00642; zf-CCCH; 1.
KW Nuclear protein; RNA-binding; mRNA splicing; Zinc-finger; Repeat.
FT DOMAIN 47 >82 INVOLVED IN BINDING U2AF-65 (BY
FT SIMILARITY).
FT NON_TER 82 82
SQ SEQUENCE 82 AA; 9416 MW; 96882458044823A8 CRC64;

Query Match 68.98; Score 31; DB 1; Length 82;
Best Local Similarity 71.48; Pred. No. 14;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NATFYFK 8

|||||

Db 17 NCSFYFK 23

RESULT 8

ID HIS5_BUCAI STANDARD; PRT; 196 AA.
AC P57204;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Amidotransferase hish (EC 2.4.2.-).
GN HISH OR BUL03.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS symbiotic bacterium).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TOKYO 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RT Buchnera sp. APS.";
RL Nature 407:81-86(2000).
CC -1- FUNCTION: CATALYZES AN AMIDOTRANSFERASE REACTION THAT GENERATES
CC IMIDAZOLE-GLYCEROL PHOSPHATE AND 5-AMINOIMIDAZOL-4-CARBOXYAMIDE
CC RIBONUCLEOTIDE, WHICH IS USED FOR PURINE SYNTHESIS.
CC -1- PATHWAY: FIFTH STEP IN HISTIDINE BIOSYNTHETIC PATHWAY.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE HISH FAMILY.
CC -1- SIMILARITY: CONTAINS 1 TYPE-1 GLUTAMINE AMIDOTRANSFERASE DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AP001118; BAB12822.1; -.
DR InterPro; IPR000991; GATase_1.
DR Pfam; PF00117; GATase_1.
DR PROSITE; PS00442; GATASE_TYPE_I; FALSE_NEG.

KW Histidine biosynthesis; Transferase; Glutamine amidotransferase;
KW Complete proteome.
FT ACT_SITE 77 77 BY SIMILARITY.
FT ACT_SITE 178 178 BY SIMILARITY.
FT ACT_SITE 180 180 BY SIMILARITY.
SQ SEQUENCE 196 AA; 21679 MW; 904624B74BBE5D CRC64;

Query Match 68.98; Score 31; DB 1; Length 196;
Best Local Similarity 71.48; Pred. No. 34;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NNATFYF 7

Db 135 NNSRFF 141
||: |||

RESULT 9

U2AG_MOUSE STANDARD; PRT; 239 AA.
AC Q9D883; Q9CZ98; Q99LX2;
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Splicing factor U2AF 35 kDa subunit (U2 auxiliary factor 35 kDa subunit) (U2 snRNP auxiliary factor small subunit).
GN U2AF1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Small intestine;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gliss C., King B., Kochiwa H., Kuhl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE OF 2-239 FROM N.A.
RX TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PLAYS A CRITICAL ROLE IN BOTH CONSTITUTIVE AND ENHANCER-DEPENDENT SPLICING BY MEDIATING ESSENTIAL PROTEIN-PROTEIN INTERACTIONS AND PROTEIN-RNA INTERACTIONS REQUIRED FOR ACCURATE 3' SPLICING SITE SELECTION. DIRECTLY MEDIATES INTERACTIONS BETWEEN U2AF65 (LARGE SUBUNIT) AND PROTEINS BOUND TO THE ENHANCER AND THIS MAY FUNCTION AS A BRIDGE BETWEEN U2AF65 AND THE ENHANCER COMPLEX TO RECRUIT IT TO THE ADJACENT INTRON (By similarity).
CC -!- SUBUNIT: ASSOCIATES WITH A 65 kDa PROTEIN (By similarity).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (By similarity).
CC -!- DOMAIN: THE AMINO-TERMINAL SR-RICH DOMAIN IS REQUIRED FOR INTERACTIONS WITH SR PROTEINS AND THE SPLICING REGULATORS TRA AND TRA2, AND THE CARBOXY-TERMINAL DOMAIN IS REQUIRED FOR FORMATION OF THE U2AF35/U2AF65 HETERODIMER (By similarity).
CC -!- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -!- SIMILARITY: CONTAINS 2 C3H1-TYPE ZINC FINGERS.
CC -!- SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AK008332; BAB25609.1; -

DR EMBL; AK012849; BAB28511.1; -
DR EMBL; BC002184; AAH02184.1; -
DR MGD; MGI:98884; U2af1.
DR InterPro; IPR000504; RRM.
DR InterPro; IPR003954; RRM_1.
DR InterPro; IPR000571; Zf-CCCH.
DR Pfam; PF00076; rrm; 1.
DR Pfam; PF0642; zf-CCCH; 2.
DR SMART; SM00360; RRM; 1.
DR SMART; SM00361; RRM_1; 1.
DR SMART; SM00356; Znf_C3H1; 2.
DR PROSITE; PS0102; RRM; 1.
DR PROSITE; PS00303; RRM_RNP_1; FALSE NEG.
KW Nuclear protein; RNA-binding; mRNA splicing; Zinc-finger; Repeat.
FT ZN_FING 13 41 C3H1-TYPE 1.
FT DOMAIN 65 147 RNA-BINDING (RRM).
FT ZN_FING 149 173 C3H1-TYPE 2.
FT DOMAIN 179 238 ARG/GLY/SER-RICH (RS DOMAIN).
FT CONFLICT 187 187 G -> R (IN REF. 1; BAB25609).
SQ SEQUENCE 239 AA; 27815 MW; DFF944210581244D CRC64;

Query Match 68.9%; Score 31; DB 1; Length 239;
Best Local Similarity 71.4%; Pred. No. 41;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 NATFYFK 8
DB 17 NCSFYFK 23
| : ||||
| : ||||

RESULT 10
U2AG_HUMAN STANDARD; PRT; 240 AA.
AC Q01081;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Splicing factor U2AF 35 kDa subunit (U2 auxiliary factor 35 kDa subunit) (U2 snRNP auxiliary factor small subunit).
GN U2AF1 OR U2AF35.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 68-89 AND 126-151.
RX TISSUE=Fetal brain;
RX MEDLINE=92409598; PubMed=1388271;
RA Zhang M., Zamore P.D., Carmo-Fonseca M., Lamond A.I., Green M.R.;
RT "Cloning and intracellular localization of the U2 small nuclear ribonucleoprotein auxiliary factor small subunit."
RL Proc. Natl. Acad. Sci. U.S.A. 89:8769-8773(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T., Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D., Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W., Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S., Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E., Minoshima S., Shimizu N., Nordiek G., Hornischer K., Brandt P., Schaefer M., Schoen O., Desario A., Hennig S., Kauer G., Bloecker H., Wehmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F., Lehrach H., Reinhardt R., Vaspo M.-L.;
RT "The DNA sequence of human chromosome 21."
RL Nature 405:311-319(2000).
RN [3]
RP FUNCTION.
RX MEDLINE=96249383; PubMed=8647433;

RA Zuo P., Maniatis T.;
 RT "The splicing factor U2AF35 mediates critical protein-protein
 interactions in constitutive and enhancer-dependent splicing."
 RL Genes Dev. 10:1356-1368(1996).
 CC -!- FUNCTION: PLAYS A CRITICAL ROLE IN BOTH CONSTITUTIVE AND ENHANCER-
 CC DEPENDENT SPLICING BY MEDIATING ESSENTIAL PROTEIN-PROTEIN
 CC INTERACTIONS AND PROTEIN-RNA INTERACTIONS REQUIRED FOR ACCURATE 3'
 CC SPLICE SITE SELECTION. DIRECTLY MEDIATES INTERACTIONS BETWEEN
 CC U2AF65 (LARGE SUBUNIT) AND PROTEINS BOUND TO THE ENHANCERS AND
 CC U2AF35 (SMALL SUBUNIT) AS A BRIDGE BETWEEN U2AF65 AND THE ENHANCER
 CC COMPLEX TO RECRUIT IT TO THE ADJACENT INTRON.
 CC -!- SUBUNIT: ASSOCIATES WITH A 65 kda PROTEIN.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- DOMAIN: THE AMINO-TERMINAL SR-RICH DOMAIN IS REQUIRED FOR
 CC INTERACTIONS WITH SR PROTEINS AND THE SPLICING REGULATORS TRA AND
 CC TRA2, AND THE CARBOXY-TERMINAL DOMAIN IS REQUIRED FOR FORMATION OF
 CC THE U2AF35/U2AF65 HETERODIMER.
 CC -!- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC -!- SIMILARITY: CONTAINS 2 C3H1-TYPE ZINC FINGERS.
 CC -!- SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M96982; AAA36619.1; -;
 DR EMBL; AP001748; BAA95534.1; -;
 DR PIR; S27969; S27969.
 DR PIR; A46179; A46179.
 DR MIM; 191317; -;
 DR InterPro: IPR000504; RRM.
 DR InterPro: IPR000571; Zf-CCCH.
 DR Pfam; PF00076; rrm; 1.
 DR Pfam; PF00642; zf-CCCH; 2.
 DR SMART; SM00360; RRM; 1.
 DR SMART; SM00356; Znf_C3H1; 2.
 DR PROSITE; PS0102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
 KW Nuclear protein; RNA-binding; mRNA splicing; zinc-finger; Repeat.
 FT ZN-FING 13 41
 FT C3H1-TYPE 1.
 FT RNA-BINDING (RRM).
 FT ZN-FING 65 147
 FT C3H1-TYPE 2.
 FT DOMAIN 149 173
 FT ARG/GLY/SER-RICH (RS DOMAIN).
 FT DOMAIN 178 240
 FT POLY-GLY
 FT DOMAIN 210 223
 SQ SEQUENCE 240 AA; 27872 MW; 3DA130DCE0B953F6 CRC64;

Query Match 68.9%; Score 31; DB 1; Length 240;
 Best Local Similarity 71.4%; Pred. No. 42;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NATFYFK 8
 | :|||
 Db 17 NCSFYFK 23

RESULT 11

U2AG_DROME STANDARD; PRT; 264 AA.
 AC Q94535; Q9VFN4;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Splicing factor U2af 38 kda subunit (U2 auxiliary factor 38 kda
 DE subunit) (U2 SNRP auxiliary factor small subunit).
 GN U2AF38 OR CG3582.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.
 RN NCBI_TaxID=7227;
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Embryo;
 RX MEDLINE=96413646; PubMed=8816800;
 RA Rudner D.Z., Kanaar R., Breger K.S., Rio D.C.;
 RT "Mutations in the small subunit of the Drosophila U2AF splicing
 RT factor cause lethality and developmental defects";
 RL proc. Natl. Acad. Sci. U.S.A. 93:10333-10337(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA Sutton G.G., Wortman J.R., Richards S.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Beeson K.Y., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
 RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 CC -!- FUNCTION: NECESSARY FOR THE SPLICING OF PRE-MRNA. BINDS TO THE
 CC POLYPYRIMIDINE TRACT OF INTRONS EARLY DURING SPLICESOME ASSEMBLY
 CC (BY SIMILARITY).
 CC -!- SUBUNIT: ASSOCIATES WITH A 65 kda PROTEIN (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC -!- SIMILARITY: CONTAINS 2 C3H1-TYPE ZINC FINGERS.
 CC -!- SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U67066; AAB17271.1; -;
 DR EMBL; AE003590; AAP51512.1; -;
 DR FlyBase; FBgn0017457; U2af38.
 DR InterPro; IPR000504; RRM.

```

DR InterPro: IPR003954; RRM_1.
DR InterPro: IPR000571; Zf-CCCH.
DR Pfam: PF00076; rrm; 1.
DR Pfam: PF00842; zf-CCCH; 2.
DR SMART: SM00361; RRM_1; 1.
DR SMART: SM00356; Znf_C3H1; 2.
DR PROSITE: PS50102; RRM; 1.
DR PROSITE: PS50030; RRM_RNP_1; FALSE_NEG.
KW Nuclear protein; RNA-binding; mRNA splicing; Zinc-finger; Repeat.
FT DOMAIN 44 149 RNA-BINDING (RRM).
FT DOMAIN 180 213 ARG/SER-RICH (RS DOMAIN).
FT DOMAIN 190 197 POLY-ARG.
FT DOMAIN 252 262 POLY-GLY.
FT CONFLICT 66 66 H -> D (IN REF. 1).
SQ SEQUENCE 264 AA; 29877 MW; 577285FB66FDB2F5 CRC64;

Query Match 68.9%; Score 31; DB 1; Length 264;
Best Local Similarity 71.4%; Pred. No. 46;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 NATFYFK 8
DB 17 NCSEYFK 23

RESULT 12
ARGC_ECOLI STANDARD; PRT; 334 AA.
AC P11446;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38) (AGPR) (N-
DE acetyl-glutamate semialdehyde dehydrogenase) (NAGSA dehydrogenase).
GN ARGC OR B3958.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=89121510; PubMed=2851495;
RA Parsot C., Boyen A., Cohen G.N., Glandsdorff N.;
RT "Nucleotide sequence of Escherichia coli argB and argC genes:
RT Comparison of N-acetylglutamate kinase and
RT N-acetylglutamate-gamma-semialdehyde dehydrogenase with homologous
RT and analogous enzymes.";
RL Gene 68:275-283(1988).
[2]
RN SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RX MEDLINE=94089392; PubMed=8265357;
RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
RA Daniels D.L.;
RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the
RT region from 89.2 to 92.8 minutes.";
RL Nucleic Acids Res. 21:5408-5417(1993).
[3]
RN SEQUENCE OF 1-48 FROM N.A.
RX STRAIN=K12;
RX MEDLINE=83143275; PubMed=6761650;
RA Plette J., Cunin R., Boyen A., Charlier D.R.M., Crabeel M.,
RA van Vliet F., Glandsdorff N., Squires C., Squires C.L.;
RT "The regulatory region of the divergent argECBH operon in Escherichia
RT coli K-12.";
RL Nucleic Acids Res. 10:8031-8048(1982).
[4]
RN SEQUENCE OF 1-19 FROM N.A.
RX STRAIN=K12;
RX MEDLINE=92202162; PubMed=1551850;
RA Meinel T., Schmitt E., Mechulam Y., Blanquet S.;
RT "Structural and biochemical characterization of the Escherichia coli

```

```

RT argE gene product.";
RL J. Bacteriol. 174:2323-2331(1992).
CC -|- CATALYTIC ACTIVITY: N-acetyl-L-glutamate 5-semialdehyde + NADP(+)
CC + phosphate = N-acetyl-5-glutamyl phosphate + NADPH.
CC -|- PATHWAY: THIRD STEP IN ARGININE BIOSYNTHESIS.
CC -|- SIMILARITY: BELONGS TO THE NAGSA DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M21446; AAA23477.1; -.
DR EMBL; J01587; AAB59146.1; -.
DR EMBL; X55417; -; NOT_ANNOTATED_CDS.
DR EMBL; U00006; AAC43064.1; -.
DR EMBL; AE000470; AAC76940.1; -.
DR PIR; J07332; RDECEP.
DR EcoGene; EG10065; argC.
DR InterPro; IPR000706; AGPR_act_site.
DR InterPro; IPR000534; Semialdh_dh.
DR Pfam; PF01118; Semialdehyde_dh; 1.
DR Pfam; PF02774; Semialdehyde_dhC; 1.
DR ProDom; PD003765; AGPR_act_site; 1.
DR PROSITE; PS01224; ARGC; 1.
KW Arginine biosynthesis; Oxidoreductase; NADP; Complete proteome.
FT ACT_SITE 154 154 BY SIMILARITY.
SQ SEQUENCE 334 AA; 35952 MW; 67AC195ECE1C4789 CRC64;

Query Match 68.9%; Score 31; DB 1; Length 334;
Best Local Similarity 75.0%; Pred. No. 58;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNATFYFK 8
DB 109 NDATFYFK 116

RESULT 13
OMPU_VIBCH STANDARD; PRT; 341 AA.
AC P97085; Q9L5A3; Q9KU90;
DT 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Outer membrane protein U precursor (Porin ompU).
DE OMPU OR VC0633.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
[1]
RN SEQUENCE FROM N.A.
RX STRAIN=CLASSICAL OGAWA 395 / ATCC 39541 / SEROTYPE O1;
RX MEDLINE=97101069; PubMed=8945596;
RA Sperandio V., Bailey C.C., Giron J.A., Dirita V.J., Silveira W.D.,
RA Vettore A.L., Kaper J.B.;
RT "Cloning and characterization of the gene encoding the OmpU outer
RT membrane protein of Vibrio cholerae.";
RL Infect. Immun. 64:5406-5409(1996).
[2]
RN SEQUENCE FROM N.A.
RX STRAIN=EL TOR;
RA Yin Y., Zhang J.Z.;
RT "Cloning and expression of ompU of Vibrio cholerae and its
RT antigenicity analysis.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RX STRAIN=EL TOR N16961 / SEROTYPE O1;

```

```

RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermlaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Ueberback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U73751; AAB48973.1; -.
CC EMBL: AF253529; AAF64526.1; -.
CC EMBL: AEO04149; AAF93799.1; ALT_INIT.
CC TIGR: VC0633; -.
CC InterPro: IPR001702; Gram_neg_porin.
CC InterPro: IPR003229; OMP_2.
CC Pfam: PF00267; Omp-ve_porins; 1.
CC ProDom: PD000808; Omp_2; 1.
CC Transmembrane; Porin; Signal; Outer membrane; Complete proteome.
KW SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 341 OUTER MEMBRANE PROTEIN U.
FT CONFLICT 278 278 F -> I (IN REF. 2).
FT CONFLICT 290 290 E -> K (IN REF. 1).
FT CONFLICT 324 325 VG -> AS (IN REF. 1).
SQ SEQUENCE 341 AA; 36645 MW; CECB39070E41732 CRC64;

Query Match 68.9%; Score 31; DB 1; Length 341;
Best Local Similarity 71.4%; Pred. No. 59;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 2 NATFYFK 8
Db 299 DATYFK 305
:|:|:|

RESULT 14
KNL2_BOVIN ID KNL2_BOVIN STANDARD; PRT; 434 AA.
AC P01047;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Kininogen, LMW II precursor (Thiol proteinase inhibitor) [Contains:
DE Bradykinin].
OS Bos taurus (Bovine).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=83117859; PubMed=6572010;
RA Nawa H., Kitamura N., Hirose T., Asai M., Inayama S., Nakanishi S.;
RT "Primary structures of bovine liver low molecular weight kininogen
RT precursors and their two mRNAs."
RL Proc. Natl. Acad. Sci. U.S.A. 80:90-94(1983).
RN (2)
RP SEQUENCE OF 19-376.
RX MEDLINE=87137530; PubMed=3546295;
RA Sueyoshi T., Miyata T., Hashimoto N., Kato H., Hayashida H.,
RA Miyata T., Iwanaga S.;

```

```

RT "Bovine high molecular weight kininogen. The amino acid sequence,
RT positions of carbohydrate chains and disulfide bridges in the heavy
RT chain portion."
RL J. Biol. Chem. 262:2768-2779(1987).
CC -1- FUNCTION: (1) KININOGENS ARE INHIBITORS OF THIOL PROTEASES; (2)
CC LMW-KININOGEN INHIBITS THE AGGREGATION OF THROMBOCYTES; (3) THE
CC ACTIVE PEPTIDE KALLIDIN THAT IS RELEASED FROM LMW-KININOGEN SHOWS
CC A VARIETY OF PHYSIOLOGICAL EFFECTS; (3A) INFLUENCE IN SMOOTH
CC MUSCLE CONTRACTION, (3B) INDUCTION OF HYPOTENSION, (3C)
CC NATRIURESIS AND DIURESIS (KIDNEY).
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- ALTERNATIVE PRODUCTS: HMW II AND LMW II KININOGEN PRECURSORS ARE
CC PRODUCED FROM THE SAME GENE AS THE RESULT OF ALTERNATE MRNA
CC SPLICING. THE SEQUENCES OF BOTH KININOGENS ARE IDENTICAL UP
CC TO RESIDUE 398.
CC -1- TISSUE SPECIFICITY: PLASMA.
CC -1- PTM: BRADYKININ IS RELEASED FROM KININOGEN BY PLASMA KALLIKREIN.
CC -1- MISCELLANEOUS: LMW-KININOGEN IS IN CONTRAST TO HMW-KININOGEN NOT
CC INVOLVED IN BLOOD CLOTTING.
CC -1- SIMILARITY: CONTAINS 3 CYSTATIN-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: V00427; CAA23710.1; -.
CC PIR: A01284; KGBOL2.
CC HSP: P01038; IA90.
CC InterPro: IPR000010; Cystatin.
CC InterPro: IPR003243; Cystatin_C_M.
CC Pfam: PF00031; cystatin; 3.
CC ProDom: PD001231; Cystatin_C_M; 1.
CC SMART: SM00043; CY; 3.
CC PROSITE: PS00287; CYSTATIN; 2.
KW Glycoprotein; Plasma; Repeat; Vasodilator; Alternative splicing;
KW Thiol protease inhibitor; Bradykinin; Signal.
FT SIGNAL 1 18
FT CHAIN 19 434 KININOGEN, LMW II.
FT CHAIN 19 376 HEAVY CHAIN.
FT PEPTIDE 378 386 BRADYKININ.
FT CHAIN 387 434 LIGHT CHAIN.
FT DOMAIN 19 135 CYSTATIN-LIKE 1.
FT DOMAIN 136 256 CYSTATIN-LIKE 2.
FT DOMAIN 257 376 CYSTATIN-LIKE 3.
FT MOD_RES 19 19 PYRROLIDONE CARBOXYLIC ACID.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .).
FT CARBOHYD 136 136 O-LINKED (PARTIAL).
FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .); OR 169,
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (PARTIAL).
FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .).
FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .).
FT DISULFID 27 404 INTERCHAIN.
FT DISULFID 82 93
FT DISULFID 106 125
FT DISULFID 141 144
FT DISULFID 205 217
FT DISULFID 228 247
FT DISULFID 261 264
FT DISULFID 325 337
FT DISULFID 348 367
SQ SEQUENCE 434 AA; 48148 MW; 73A7079DE3E03430 CRC64;

Query Match 68.9%; Score 31; DB 1; Length 434;
Best Local Similarity 62.5%; Pred. No. 76;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 NNATFYFK 8
::|:|:|:|

```


Db 290 HDGTFYFK 297

RESULT 15

PAFA_HUMAN

ID PAFA_HUMAN STANDARD; PRT; 441 AA.

AC Q13093; Q15692;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Platelet-activating factor acetylhydrolase precursor (EC 3.1.1.47)

DE (PAF acetylhydrolase) (PAF 2-acetylhydrolase) (LPL-associated

DE phospholipase A2) (LPL-PLA2)) (2-acetyl-1-alkylglycerophosphocholine

DE esterase) (1-alkyl-2-acetyl-glycerophosphocholine esterase).

GN PLA2G7 OR PAFAH.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 42-57.

RC TISSUE=Myeloid;

RX MEDLINE=95214779; PubMed=7700381;

RA Tjoelker L.W., Wilder C., Eberhardt C., Stafforini D.M., Dietsch G.,

RA Schimpf B., Hooper S., le Trong H., Cousins L.S., Zimmerman G.A.,

RA Yamada Y., McIntyre T.M., Prescott S.M., Gray P.W.;

RT "Anti-inflammatory properties of a platelet-activating factor

RT acetylhydrolase.";

RL Nature 374:549-553(1995).

RN [2]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RC TISSUE=Lymphoma;

RX MEDLINE=96197208; PubMed=8624782;

RA Tew D.G., Southern C., Rice S.Q.J., Lawrence M.P., Li H., Boyd H.F.,

RA Moores K., Gloger I.S., Macphee C.H.;

RT "Purification, properties, sequencing, and cloning of a lipoprotein-

RT associated, serine-dependent phospholipase involved in the oxidative

RT modification of low-density lipoproteins.";

RL Arterioscler. Thromb. Vasc. Biol. 16:591-599(1996).

RN [3]

RP MUTAGENESIS.

RX MEDLINE=96029630; PubMed=7592717;

RA Tjoelker L.W., Eberhardt C., Unger J., le Trong H.,

RA Zimmerman G.A., McIntyre T.M., Stafforini D.M., Prescott S.M.,

RA Gray P.W.;

RT "Plasma platelet-activating factor acetylhydrolase is a secreted

RT phospholipase A2 with a catalytic triad.";

RL J. Biol. Chem. 270:25481-25487(1995).

RN [4]

RP VARIANT PHE-279.

RX MEDLINE=96259525; PubMed=8675689;

RA Stafforini D.M., Satoh K., Atkinson D.L., Tjoelker L.W.,

RA Eberhardt C., Yoshida H., Imaizumi T., Takamatsu S., Zimmerman G.A.,

RA McIntyre T.M., Gray P.W., Prescott S.M.;

RT "Platelet-activating factor acetylhydrolase deficiency. A missense

RT mutation near the active site of an anti-inflammatory

RT phospholipase.";

RL J. Clin. Invest. 97:2784-2791(1996).

RN [5]

RP VARIANT PLA2G7 DEFICIENCY ARG-281.

RX MEDLINE=97396177; PubMed=9245731;

RA Yamada Y., Yokota M.;

RT "Loss of activity of plasma platelet-activating factor acetylhydrolase

RT due to a novel Gln281-->Arg mutation.";

RL Biochem. Biophys. Res. Commun. 236:772-775(1997).

RN [6]

RP VARIANT PHE-279.

RX MEDLINE=98074100; PubMed=9412624;

RA Hiramoto M., Yoshida H., Imaizumi T., Yoshimizu N., Satoh K.;

RT "A mutation in plasma platelet-activating factor acetylhydrolase

RT (Val279-->Phe) is a genetic risk factor for stroke.";

RL Stroke 28:2417-2420(1997).

RN [7]

RP VARIANT PHE-279.

RX MEDLINE=98132308; PubMed=9472966;

RA Yamada Y., Ichihara S., Fujimura T., Yokota M.;

RT "Identification of the G994-->T missense in exon 9 of the plasma

RT platelet-activating factor acetylhydrolase gene as an independent

RT risk factor for coronary artery disease in Japanese men.";

RL Metabolism 47:177-181(1998).

RN [8]

RP VARIANT PHE-279.

RX MEDLINE=98430412; PubMed=9759612;

RA Yoshida H., Imaizumi T., Fujimoto K., Itaya H., Hiramoto M.,

RA Yoshimizu N., Fukushi K., Satoh K.;

RT "A mutation in plasma platelet-activating factor acetylhydrolase

RT (Val279phe) is a genetic risk factor for cerebral hemorrhage but not

RT for hypertension.";

RL Thromb. Haemost. 80:372-375(1998).

RN [9]

RP VARIANTS HIS-92; THR-198 AND ALA-379.

RX MEDLINE=20311534; PubMed=10733466;

RA Kruse S., Mao X.-Q., Heinzmann A., Blattmann S., Roberts M.H.,

RA Braun S., Gao P.-S., Forster J., Kuehr J., Hopkin J.M., Shirakawa T.,

RA Deichmann K.A.;

RT "The Ile198Thr and Ala379Val variants of plasmatic PAF-acetylhydrolase

RT impair catalytical activities and are associated with atopy and

RT asthma.";

RL Am. J. Hum. Genet. 66:1522-1530(2000).

CC -!- FUNCTION: MODULATES THE ACTION OF PLATELET-ACTIVATING FACTOR (PAF)

CC BY HYDROLYZING THE SN-2 ESTER BOND TO YIELD THE BIOLOGICALLY

CC INACTIVE LYSO-PAF. HAS A SPECIFICITY FOR SUBSTRATES WITH A SHORT

CC RESIDUE AT THE SN-2 POSITION. IT IS INACTIVE AGAINST LONG-CHAIN

CC PHOSPHOLIPIDS.

CC -!- CATALYTIC ACTIVITY: 2-acetyl-1-alkyl-sn-glycero-3-phosphocholine +

CC H(2)O = 1-alkyl-sn-glycero-3-phosphocholine + acetate.

CC -!- SUBCELLULAR LOCATION: Extracellular.

CC -!- TISSUE SPECIFICITY: PLASMA.

CC -!- POLYMORPHISM: THE POLYMORPHIC ALLELES THR-198 AND VAL-379 ARE

CC ASSOCIATED WITH ATOPY AND ASTHMA.

CC -!- DISEASE: DEFECTS IN PLA2G7 ARE THE CAUSE OF PLA2G7 DEFICIENCY A

CC TRAIT WHICH IS PRESENT IN 2% OF JAPANESE. IT COULD HAVE A

CC SIGNIFICANT PHYSIOLOGIC EFFECT IN THE PRESENCE OF INFLAMMATORY

CC BODILY RESPONSES.

CC -!- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,

CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; U20157; AAC50126.1; -

CC EMBL; U24577; AAB04170.1; -

CC MIM; 601690; -

CC InterPro: IPR000379; Est_lip_thioest_actsite.

CC InterPro: IPR000734; Lipase.

CC PROSITE; PS00120; LIPASE_SER; 1.

CC Hydrolase; Lipid degradation; Glycoprotein; Signal; Polymorphism;

CC Disease mutation.

FT SIGNAL 1 21

FT CHAIN 22 441 PLATELET-ACTIVATING FACTOR

FT ACT_SITE 273 273 ACETYLHYDROLASE

FT ACT_SITE 296 296 CHARGE RELAY SYSTEM.

FT ACT_SITE 351 351 CHARGE RELAY SYSTEM.

FT CARBOHYD 423 423 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 433 433 N-LINKED (GLCNAC. . .)

FT VARIANT 92 92 R -> H (COMMON POLYMORPHISM; IN

FT DBSNP:1805017).

FT FTID-VAR.011583.

FT I -> T (COMMON POLYMORPHISM; IN

FT DBSNP:1805018).

FT VARIANT 198 198

```
FT FTid=VAR_011584.
FT V -> F (IN PLA2G7 DEFICIENCY; LOSS OF
FT FUNCTION; MORE COMMON AMONG JAPANESE THAN
FT IN CAUCASIANS; RISK FACTOR FOR CORONARY
FT ARTERY DISEASE AND STROKE).
FT /FTid=VAR_004268.
FT Q -> R (IN PLA2G7 DEFICIENCY; LOSS OF
FT FUNCTION).
FT /FTid=VAR_011585.
FT V -> A (COMMON POLYMORPHISM).
FT /FTid=VAR_011586.
FT S->A: ACTIVITY IS HIGHER THAN WILD TYPE.
FT S->A: LOSS OF ACTIVITY.
FT D->A: ALMOST NO ACTIVITY.
FT D->N: DIMINISHED ACTIVITY.
FT D->A: LOSS OF ACTIVITY.
FT D->N: LOSS OF ACTIVITY.
FT D->A: NO CHANGE IN ACTIVITY.
FT D->A: ACTIVITY IS HIGHER THAN WILD TYPE.
FT H->A: LOSS OF ACTIVITY.
SQ SEQUENCE 441 AA; 50077 MW; 3BA9EEA9E8094A57 CRC64;
```

Query Match 68.9%; Score 31; DB 1; Length 441;
Best Local Similarity 71.4%; Pred. NO. 77;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 NATFYFK 8
Db :|||
185 SATYFK 191

Search completed: July 10, 2002, 09:22:11
Job time: 97 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 10, 2002, 09:20:34 ; Search time 25.02 Seconds
(without alignments)
55.314 Million cell updates/sec

Title: US-09-461-061a-1

Perfect score: 45

Sequence: 1 NNATFYFK 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_unclassified:*
- 14: sp_vertebrate:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	36	80.0	858	5 Q21853	Q21853 caenorhabdi
2	35	77.8	182	5 Q21968	Q21968 caenorhabdi
3	34	75.6	71	16 Q9KDI7	Q9KDI7 bacillus ha
4	34	75.6	115	15 Q9YXT3	Q9YXT3 human immu
5	34	75.6	173	10 Q9ZRF8	Q9ZRF8 oryza sativ
6	34	75.6	181	10 Q9M625	Q9M625 hordeum vul
7	34	75.6	182	10 P93615	P93615 triticeum ae
8	34	75.6	273	16 Q9CD90	Q9CD90 mycobacteri
9	34	75.6	319	16 Q9CJ05	Q9CJ05 lactococcus
10	34	75.6	326	12 Q66771	Q66771 equine rota
11	34	75.6	326	12 Q66772	Q66772 equine rota
12	34	75.6	326	12 Q9IE87	Q9IE87 human rotav
13	34	75.6	357	13 Q91655	Q91655 xenopus lae
14	34	75.6	440	8 Q9ZY23	Q9ZY23 pedinomonas
15	34	75.6	497	10 Q9FVS9	Q9FVS9 arabidopsis
16	34	75.6	602	2 Q9AFA2	Q9AFA2 staphylococ

17	34	75.6	630	5 P91391	P91391 caenorhabdi
18	34	75.6	758	5 O16610	O16610 caenorhabdi
19	34	75.6	949	16 Q97DM1	Q97DM1 clostridium
20	33	73.3	251	10 Q9LDL9	Q9LDL9 arabidopsis
21	33	73.3	303	5 O44623	O44623 caenorhabdi
22	33	73.3	332	10 O64597	O64597 arabidopsis
23	33	73.3	469	17 Q29037	Q29037 archaeglob
24	33	73.3	706	10 Q9S7L6	Q9S7L6 arabidopsis
25	33	73.3	1346	5 O45699	O45699 caenorhabdi
26	32	71.1	155	8 Q79080	Q79080 lipolexis g
27	32	71.1	331	12 Q91AP3	Q91AP3 human rotav
28	32	71.1	537	10 Q9LH23	Q9LH23 oryza sativ
29	32	71.1	606	8 Q9GLW4	Q9GLW4 sus scrofa
30	32	71.1	606	8 Q9TDR1	Q9TDR1 sus scrofa
31	32	71.1	606	8 Q99997	Q99997 sus scrofa
32	32	71.1	606	8 Q9G7T6	Q9G7T6 sus scrofa
33	32	71.1	606	8 Q9G7R8	Q9G7R8 sus scrofa
34	32	71.1	630	12 Q9EML1	Q9EML1 amsacta moo
35	32	71.1	1096	4 Q94836	Q94836 homo sapien
36	32	71.1	1191	5 Q9VJN7	Q9VJN7 drosophila
37	32	71.1	1215	5 Q22649	Q22649 caenorhabdi
38	32	71.1	1255	5 Q9NK83	Q9NK83 drosophila
39	32	71.1	1560	5 Q9GRX5	Q9GRX5 dictyosteli
40	32	71.1	1561	5 Q9U987	Q9U987 dictyosteli
41	32	71.1	1731	5 Q95W43	Q95W43 trypanosoma
42	31	68.9	49	13 Q98TR0	Q98TR0 gallus gall
43	31	68.9	110	11 Q9JLF2	Q9JLF2 rattus norv
44	31	68.9	118	4 Q9UI66	Q9UI66 homo sapien
45	31	68.9	121	15 Q9Q7U5	Q9Q7U5 human immu

ALIGNMENTS

RESULT 1

Q21853 PRELIMINARY; PRT; 858 AA.
 ID Q21853;
 AC Q21853;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE R09A8.2 PROTEIN.
 GN R09A8.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wilkinson J.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL; Z68009; CAA92004.1; -;
 SQ SEQUENCE 858 AA; 96262 MW; 60E1070CA8F4D4C8 CRC64;

Query Match 80.0%; Score 36; DB 5; Length 858;
 Best Local Similarity 75.0%; Pred. No. 89;
 Matches. 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNATFYFK 8

Db 49 NESTFYFK 56

RESULT 2

Q21968

ID Q21968 PRELIMINARY; PRT; 182 AA.
 AC Q21968;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE R12H7.3 PROTEIN.
 GN R12H7.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Coles L.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology.";
 RL Science 282:2012-2018(1998).
 DR EMBL; Z50755; CAA90635.1; -;
 DR InterPro; IPR001232; Skp1.
 DR Pfam; PF01466; Skp1; 1.
 DR SMART; SMO0512; Skp1; 1.
 SQ SEQUENCE 182 AA; 21178 MW; 4BE38A19C4FA8124 CRC64;

Query Match 77.8%; Score 35; DB 5; Length 182;
 Best Local Similarity 75.0%; Pred. No. 31;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNATFFYK 8
 |||||:
 Db 161 NNATLFFK 168

RESULT 3
 ID Q9KDI7 PRELIMINARY; PRT; 71 AA.
 AC Q9KDI7;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE BH1226 PROTEIN.
 GN BH1226.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fuji F., Hirata C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis.";
 RL Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL; AP001511; BAB04945.1; -;
 KW Complete proteome.
 SQ SEQUENCE 71 AA; 8026 MW; 311AC9AEB3C539D3 CRC64;

Query Match 75.6%; Score 34; DB 16; Length 71;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NATFFYF 7
 |||||
 Db 43 NATFFYF 48

RESULT 4
 ID Q9YXT3 PRELIMINARY; PRT; 115 AA.
 AC Q9YXT3;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE ENVELOPE GLYCOPROTEIN C2V3 REGION (FRAGMENT).
 GN ENV.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RJ96BRP071;
 RA Tanuri A., Swanson P.A., Devare S.G., Berro O.J., Savedra A.,
 RA Costa L.J., Telles J.G., Brindeiro R., Schable C., Pieniazek D.,
 RA Rayfield M.;
 RT "HIV-1 subtypes among blood donors from Rio de Janeiro, Brazil.";
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF034019; AAC79271.1; -;
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 KW AIDS; Coat protein; Glycoprotein.
 FT NON_TER 1
 FT NON_TER 115
 SQ SEQUENCE 115 AA; 12755 MW; 6551E67B32DCF56C CRC64;

Query Match 75.6%; Score 34; DB 15; Length 115;
 Best Local Similarity 62.5%; Pred. No. 31;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNATFFYK 8
 |||||:
 Db 78 NNTTFFFR 85

RESULT 5
 ID Q9ZRF8 PRELIMINARY; PRT; 173 AA.
 AC Q9ZRF8;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE HYDROPHOBIC LEA-LIKE PROTEIN.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=L0MELLO;
 RA Chen L.J., Chai Y.J., Chen P.W.;
 RT "A rice embryo-specific gene with high homology to soybean GmPM3 gene,
 RT a hydrophobic LEA protein gene.";
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U57639; AAD10377.1; -;
 SQ SEQUENCE 173 AA; 18287 MW; 63FA2F778BB7259D CRC64;

Query Match 75.6%; Score 34; DB 10; Length 173;
 Best Local Similarity 85.7%; Pred. No. 46;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNATFFYF 7
 |||||
 Db 47 NGATFFYF 53

```

RESULT 6
ID Q9M625 PRELIMINARY; PRT; 181 AA.
AC Q9M625;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE PLASMA MEMBRANE ASSOCIATED PROTEIN.
GN PM19.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RA Morris P.C., Ranford J.C.;
RT "Hordeum vulgare hydrophobic embryo-associated protein PM19 mRNA.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF218627; AAF29532.1; -.
SQ SEQUENCE 181 AA; 19036 MW; 849727F1123A4030 CRC64;

Query Match 75.6%; Score 34; DB 10; Length 181;
Best Local Similarity 85.7%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNATYFF 7
DB 47 NGATYFF 53

RESULT 7
ID P93615 PRELIMINARY; PRT; 182 AA.
AC P93615;
DT 01-MAY-1997 (TRENBLrel. 03, Created)
DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE ABA INDUCED PLASMA MEMBRANE PROTEIN PM 19.
GN WTABAPW.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. CHIHOKU;
RX MEDLINE=97393494; PubMed=9249988;
RA Koike M., Takezawa D., Arakawa K., Yoshida S.;
RT "Accumulation of 19-kDa plasma membrane polypeptide during induction
of freezing tolerance in wheat suspension-cultured cells by abscisic
acid.";
RL Plant Cell Physiol. 38:707-716(1997).
DR EMBL; U80037; AAB38504.1; -.
SQ SEQUENCE 182 AA; 19009 MW; 799B99994AEA87DCB CRC64;

Query Match 75.6%; Score 34; DB 10; Length 182;
Best Local Similarity 85.7%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNATYFF 7
DB 47 NGATYFF 53

RESULT 8
ID Q9CD90 PRELIMINARY; PRT; 273 AA.
AC Q9CD90;
DT 01-JUN-2001 (TRENBLrel. 17, Created)

```

```

DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TRENBLrel. 18, Last annotation update)
DE HYPOTHETICAL PROTEIN ML0126.
GN ML0126.
OS Mycobacterium leprae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Horsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajadream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
DR EMBL; AL583917; CAC29634.1; -.
DR Leproma; ML0126; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 273 AA; 30831 MW; D5B765A65EF39549 CRC64;

Query Match 75.6%; Score 34; DB 16; Length 273;
Best Local Similarity 85.7%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNATYFF 7
DB 20 NNAQYFF 26

RESULT 9
ID Q9CJ05 PRELIMINARY; PRT; 319 AA.
AC Q9CJ05;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE RHANOSYLTRANSFERASE.
GN RGPB.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
lactis ssp. lactis IL1403.";
RL Genome Res. 11:731-753(2001).
DR EMBL; AE006258; AAK04299.1; -.
DR HSP; P39621; ICG8.
DR InterPro; IPR001173; Glycos_transf_2.
DR Pfam; PF00535; Glycos_transf_2; 1.
KW Transferase; Complete proteome.
SQ SEQUENCE 319 AA; 37340 MW; DIDAC78AB0A950CA CRC64;

Query Match 75.6%; Score 34; DB 16; Length 319;
Best Local Similarity 85.7%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNATYFF 7

```

Db 80 NNADEYF 86
||| |||

RESULT 10

Q66771 ID Q66771 PRELIMINARY; PRT; 326 AA.
AC Q66771;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE GLYCOPROTEIN VP7.
OS Equine rotavirus.
OC Viruses; dsRNA viruses; Reoviridae; Rotavirus.
OX NCBI_TaxID=10937;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FR4;
RX MEDLINE=95113988; PubMed=7814511;
RA Charlet M., Reggati F., Pina C.I., Liprandi F.,
RT "Equine rotaviruses with G14 serotype specificity circulate among
RT Venezuelan horses"; 32:2609-2612(1994).
RL J. Clin. Microbiol. 32:2609-2612(1994).
DR EMBL; U05348; AAA81914.1; -;
DR InterPro; IPR001963; VP7.
DR Pfam; PF00434; VP7; 1.
SQ SEQUENCE 326 AA; 37296 MW; 9FC1BD4F4CC76529 CRC64;

Query Match 75.6%; Score 34; DB 12; Length 326;
Best Local Similarity 62.5%; Pred. No. 84;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNADEYF 8
|:|||||:

Db 318 NSATFYR 325

RESULT 11

Q66772 ID Q66772 PRELIMINARY; PRT; 326 AA.
AC Q66772;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE GLYCOPROTEIN VP7.
OS Equine rotavirus.
OC Viruses; dsRNA viruses; Reoviridae; Rotavirus.
OX NCBI_TaxID=10937;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FR8;
RX MEDLINE=95113988; PubMed=7814511;
RA Charlet M., Reggati F., Pina C.I., Liprandi F.,
RT "Equine rotaviruses with G14 serotype specificity circulate among
RT Venezuelan horses"; 32:2609-2612(1994).
RL J. Clin. Microbiol. 32:2609-2612(1994).
DR EMBL; U05349; AAA67342.1; -;
DR InterPro; IPR001963; VP7.
DR Pfam; PF00434; VP7; 1.
SQ SEQUENCE 326 AA; 37352 MW; 9095E64B13933E29 CRC64;

Query Match 75.6%; Score 34; DB 12; Length 326;
Best Local Similarity 62.5%; Pred. No. 84;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNADEYF 8
|:|||||:

Db 318 NSATFYR 325

RESULT 12

Q91E87

ID Q91E87 PRELIMINARY; PRT; 326 AA.
AC Q91E87;
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE CAPSID PROTEIN.
GN VP7.
OS Human rotavirus.
OC Viruses; dsRNA viruses; Reoviridae; Rotavirus.
OX NCBI_TaxID=10941;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CC425, P3[9],G3;
RA Griffin D.D., Nakagomi T., Hoshino Y., Nakagomi O., Kirkwood C.D.,
RA Parashar U.D., Glass R.I., Gentsch J.R.;
RT "Characterization of nontypeable rotavirus strains from the United
RT States: identification of a new rotavirus reassortant (P2A[6],G12) and
RT rare P3[9] strains related to bovine rotaviruses.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ311738; CAC43312.1; -;
SQ SEQUENCE 326 AA; 37119 MW; 8B97ED1DBBD1C981 CRC64;

Query Match 75.6%; Score 34; DB 12; Length 326;
Best Local Similarity 62.5%; Pred. No. 84;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNADEYF 8
|:|||||:

Db 318 NSATFYR 325

RESULT 13

Q91655 ID Q91655 PRELIMINARY; PRT; 357 AA.
AC Q91655;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE GENE 17 PROTEIN.
GN GENE 17.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96312514; PubMed=8700860;
RA Brown D.D., Wang Z., Furrow J.D., Kanamori A., Schwartzman R.A.,
RA Remo B.F., Pinder A.;
RT "The thyroid hormone-induced tail resorption program during Xenopus
RT laevis metamorphosis.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:1924-1929(1996).
DR EMBL; U41860; AAC59876.1; -;
DR InterPro; IPR001507; zona_pellucida.
DR Pfam; PF00100; zona_pellucida; 1.
DR SMART; SM00241; ZP; 1.
SQ SEQUENCE 357 AA; 39090 MW; 5493352C8EFA21E6 CRC64;

Query Match 75.6%; Score 34; DB 13; Length 357;
Best Local Similarity 85.7%; Pred. No. 92;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NNADEYF 7
|:|||||:

Db 141 NNATLYF 147

RESULT 14

Q92Y23

Q92Y23 PRELIMINARY; PRT; 440 AA.
AC 01-WAY-1999 (TREMBlrel. 10, Created)
DT 01-WAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NADH DEHYDROGENASE SUBUNIT 2 (EC 1.6.5.3).
GN NAD2.
OS Pedinomonas minor.
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Chlorophyta; Pedinophyceae; Pedinomonadales;
OC Pedinomonadaceae; Pedinomonas.
OX NCBI_TaxID=3159;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-UTEX LB 1350;
RX MEDLINE-99418884; PubMed-10488238;
RA Turnell M., Lemieux C., Burger G., Lang B.F., Otis C., Plante I.,
RA Gray M.W.;
RA "The complete mitochondrial DNA sequences of Nephroselmis olivacea and
RT Pedinomonas minor: two radically different evolutionary patterns
RT within green algae.";
RL Plant Cell 11:1717-1729(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-UTEX LB 1350;
RA Burger G.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DDJ databases.
CC -I- CATALYTIC ACTIVITY: NADH + UBIQUINONE - NAD(+) + UBIQUINOL.
DR EMBL; AF116775; AAD19671.1; -;
DR InterPro; IPR001750; Oxidored_q1.
DR Pfam; PF00361; oxidored_q1; 1.
KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
SQ SEQUENCE 440 AA; 51380 MW; C6D6CBAD72549B4C CRC64;

Query Match 75.6%; Score 34; DB 8; Length 440;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NNATFFV 7
Db 238 NNATFFV 244

RESULT 15
Q9FVS9 PRELIMINARY; PRT; 497 AA.
AC Q9FVS9;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CYTOCHROME P450, PUTATIVE.
GN T8L23.21.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLOMBIA;
RX MEDLINE-21016719; PubMed-11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S., S.Y.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu X.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,

RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzer S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana";
RL Nature 408:816-820(2000).
CC -I- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AC079733; AAG50737.1; -;
DR InterPro; IPR001128; CYC_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; UNKNOWN1.
KW Heme: Hypothetical protein; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 497 AA; 57696 MW; COA37BA59164AB78 CRC64;

Query Match 75.6%; Score 34; DB 10; Length 497;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 NATFFVK 8
Db 64 NLTFVK 70

Search completed: July 10, 2002, 09:22:44
Job time: 130 sec

